

Wed Apr 23 13:51:14 2003

us-09-806-276a-2.rpr

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: April 23, 2003, 10:58:30 ; Search time 15.0943 seconds
(without alignments)
1617.710 Million cell updates/sec

Title: US-09-806-276A-2
Perfect score: 1326
Sequence: 1 MPAKTPYLYKAAANNKKKKK...QLDLGPSLDEVLNVMKDK 254

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_73:***
1: p1r1:***
2: p1r2:***
3: p1r3:***
4: p1r4:***

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1326	100.0	254	T46465	hypothetical prote
2	195	14.7	391	A42973	serum protein MSE5
3	109.5	8.3	1475	S42718	nuclear pore compl
4	105	7.9	2351	EZH1	coagulation factor
5	99.5	7.5	398	T43225	pheromone receptor
6	97	7.3	2271	F90073	hypothetical prote
7	96.5	7.3	1447	T00530	transcription fact
8	96	7.2	825	S26706	frequency clock pr
9	95.5	7.2	1015	T42013	hypothetical prote
10	95	7.2	482	T22981	hypothetical prote
11	94	7.1	598	T45793	hypothetical prote
12	93.5	7.1	2500	WMHUE2	HIV-EP2 enhancer-b
13	93.5	7.1	3122	T17202	DNA-directed DNA p
14	93.5	7.1	606	T51880	hypothetical prote
15	93	7.0	1337	T38949	hypothetical prote
16	92.5	7.0	667	A41311	transcription fact
17	92	6.9	817	T22442	hypothetical prote
18	92	6.9	1042	F91123	evolved beta-D-gal
19	92	6.9	1042	E85968	serine-rich protei
20	92	6.9	534	T39903	Acta protein - Lis
21	91.5	6.9	787	S47245	probable RNA-bindi
22	91	6.9	1080	T43164	probable coiled co
23	90.5	6.8	849	T01286	bud emergence prot
24	90.5	6.8	899	S62428	probable glycolipi
25	90.5	6.8	1128	A49960	TCOF1 protein - mo
26	89.5	6.7	510	T37541	DNA (cytosine-5-)-
27	89.5	6.7	1320	JC5630	
28	89.5	6.7	1612	JC5210	
29	89.5	6.7	1612	JC5210	

30	89	6.7	289	A34783
31	89	6.7	1072	T18802
32	89	6.7	1733	S27939
33	88.5	6.7	414	T43415
34	88.5	6.7	642	S34416
35	88	6.6	933	S41539
36	88	6.6	1744	A54970
37	87.5	6.6	972	T16094
38	87	6.6	542	T19952
39	87	6.6	1792	A57075
40	86	6.5	478	S56904
41	86	6.5	580	T37664
42	86	6.5	1027	T46481
43	86	6.5	1347	T02214
44	86	6.5	1983	T00385
45	85.5	6.4	862	T01141

ALIGNMENTS

RESULT 1
T46465
hypothetical protein DKFZP434A0530.1 - human
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46465
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T46465
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-254 <AAA>
A:Cross-references: EMBL:AL136842
A:Experimental source: adult testis; clone DKFZP434A0530
C:Genetics:
A:Note: DKFZP434A0530.1

Query Match 100.0%; Score 1326; DB 2; Length 254;
Best local similarity 100.0%; Pred. No. 6.6e-107;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MPAKTPYLYKAAANNKKKKFKLRDILSPDMISPLGDFRHTTHIGKEGQHDVFGDISFLQ	60
DB	1	MPAKTPYLYKAAANNKKKKFKLRDILSPDMISPLGDFRHTTHIGKEGQHDVFGDISFLQ	60
QY	61	GNVELLPNGNEKAHLGQFPGHNEFFRANSTSDSVFTETPSVLKNAISLPTIGSQALML	120
DB	61	GNVELLPNGNEKAHLGQFPGHNEFFRANSTSDSVFTETPSVLKNAISLPTIGSQALML	120
QY	121	PLSPYTFNSKQESFGPAKLPRLSCEPVMKEKQESSLLENGTVHOGDTSWSSGSASQ	180
DB	121	PLSPYTFNSKQESFGPAKLPRLSCEPVMKEKQESSLLENGTVHOGDTSWSSGSASQ	180
QY	181	SSQGRDSSSSLSSEQYPDMPAEDMFDPHPTCELIKGTKESESLDGLSLSLQDLGP	240
DB	181	SSQGRDSSSSLSSEQYPDMPAEDMFDPHPTCELIKGTKESESLDGLSLSLQDLGP	240
QY	241	SLIDEVLNVMKDK 254	
DB	241	SLIDEVLNVMKDK 254	

RESULT 2
A42973
serum protein MSE55 - human
C:Species: Homo sapiens (man)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A42973
R:Bahou, W.F.; Campbell, A.D.; Wicha, M.S.
J. Biol. Chem. 267, 13986-13992, 1992
A:Title: CDNA cloning and molecular characterization of MSE55, a novel human serum co

myogenesis protein
hypothetical prote
tensin - chicken
probable MADS-box
transcription fact
fibrinogen-binding
tensin, cardiac mu
hypothetical prote
hypothetical prote
tensin - chicken
hypothetical prote
hypothetical prote
ubiquitous TPR mot
KIAA0624 protein -
hypothetical prote

Query Match	14.7%	Score	195;	DB 2;	Length	391.
Best Local Similarity	30.0%	Pred	No.			
Matches	74.	Count				

```

RESULT 3
S42718
nuclear pore complex protein nup153 - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence
C:Accession: S42718; S37477
R:McMorrow, I.; Bastos, R.; Horton, H.; Burke, B.
Biochim. Biophys. Acta 1217, 219-223, 1994
A:Title: Sequence analysis of a cDNA encoding a human nuclear pore complex protein, hnp153
A:Reference number: S42718; MUID:94154002; PMID:8110839
A:Accession: S42718
A:Molecule type: mRNA
A:Residues: 1-1475 <MCM>
A:Cross-references

```

N; Alternate names: antihemophilic factor A; coagulation factor VIIIC; procoagulant coagulation factor VIIIc

C; Species: Homo sapiens (man)

C; Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000

C; Accession: I54318, A00525, I58059; A23584; A26174; A42348; A43986; S63527; S66445;

R; Gitschier, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

A; Title: Sequence of the exon containing regions of the human factor VIII gene.

A; Reference number: I54318; MUID:93265012; PMID:1303178

A; Status: preliminary; translated from GB/EMBL/DDBJ

A; Molecule type: DNA

A; Residues: 1-1921, 'S', 1923-2351 <RES>

A; Cross-references: GB:M88648; NID:g182381; PIDN:AAA52420.1; PID:g182383

R; Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg, P.H.

Nature 312, 330-337, 1984

A; Title: Expression of active human factor VIII from recombinant DNA clones.

A; Reference number: A00525; MUID:85061548; PMID:6438526

A; Accession: A00525

A; Molecule type: mRNA

A; Residues: 1-2351 <MOO>

A; Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179

R; Toole, J.U.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D.

Nature 312, 342-347, 1984

A; Title: Molecular cloning of a cDNA encoding human antihemophilic factor.

A; Reference number: I58059; MUID:85061550; PMID:6438528

A; Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ

A; Molecule type: protein

A; Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>

A; Cross-references: GB:K01740; NID:g182802; PIDN:AAA52484.1; PID:g182803

R; Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kohn, J.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.

DNA 4, 333-349, 1985

A; Title: Characterization of the polypeptide composition of human factor VIII:C and t

A; Reference number: A23584; MUID:86081164; PMID:3935400

A; Accession: A23584

A; Molecule type: mRNA

A; Residues: 1-2351 <TRU>

A; Cross-references: GB:M14113; NID:g182817; PIDN:AAA52485.1; PID:g182818

R; Eaton, D.; Rodriguez, H.; Vehar, G.A.

Biochemistry 25, 505-512, 1986

A; Title: Proteolytic processing of human factor VIII. Correlation of specific cleavage sites with biological activity.

A; Reference number: A26174; MUID:86159740; PMID:3082357

A; Accession: A26174

A; Molecule type: protein

A; Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <EA

R; Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992

A; Title: Identification and functional importance of tyrosine sulfate residues within human factor VIII

A; Reference number: A42348; MUID:92207952; PMID:1554716

A; Accession: A42348

A; Molecule type: protein

A; Residues: 20-36; 356-371; 392-408; 582-594; 1668-1669, 'X', 1671; 1672-1692; 1693-1708; 1709

A; Note: sequence extracted from NCBI backbone and corrected to correspond with the pu

R; Fay, P.J.; Smudzyn, T.M.

J. Biol. Chem. 264, 14005-14010, 1989

A; Title: Intersubunit fluorescence energy transfer in human factor VIII.

A; Reference number: A43986; MUID:89340500; PMID:2503509

A; Accession: A43986

A; Molecule type: protein

A; Residues: 'X', 517-523; 1853-1860, 'X', 1862-1864, 'X', 1866 <FAY>

R; Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, J.

J. Biol. Chem. 266, 740-746, 1991

A; Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential f

A; Reference number: A56109; MUID:91093266; PMID:1898735

R; Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar, P.J.

Nature 312, 326-330, 1984

A; Title: Characterization of the human factor VIII gene.

A; Reference number: A56196; MUID:85061547; PMID:6438525

Wed Apr 23 13:51:14 2003

us-09-806-276a-2.rpr

A:Contents: annotation: introns
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains of
 A:Reference number: A56216; MUID:95338127; PMID:7613471
 A:Contents: annotation: disulfide bonds
 A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
 R:Kjalkke, M.; Hedberg, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A:Reference number: S63527; MUID:96163459; PMID:8575434
 A:Accession: S63527
 A:Molecule type: protein
 A:Residues: 733-752;753-759 <KDA>
 R:Jind, P.; Larsson, K.; Sjöström, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A:Reference number: S66445; MUID:96048024; PMID:7556150
 A:Accession: S66445
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1668-1685 <LIN>
 C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
 C:Genetics:
 A:Gene: GDB:F8C
 A:Cross-references: GDB:119124; OMIM:306700
 A:Map position: Xq28-Xq28
 A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C:Function: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 A:Pathway: blood coagulation
 A:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:1-19/Domains: signal sequence #status predicted <ACH>
 F:20-2351/Product: coagulation factor VIIa heavy chain #status experimental
 F:20-740/Product: coagulation factor VIIa heavy chain #status experimental
 F:20-356/Domains: A1 <DA1>
 F:20-356/Domains: ferroxidase repeat homology <FO1>
 F:23-348/Domains: A2 <DA2>
 F:392-759/Domains: ferroxidase repeat homology <FO2>
 F:402-730/Domains: B <DB0>
 F:760-1667/Domains: coagulation factor VIIa light chain #status experimental <ACL>
 F:1668-2351/Product: coagulation factor VIIa light chain #status experimental
 F:1709-2038/Domains: A3 <DA3>
 F:1716-2038/Domains: ferroxidase repeat homology <FO3>
 F:2039-2191/Domains: C1 <DC1>
 F:2039-2188/Domains: C2 <DC2>
 F:2039-2188/Domains: I amino-terminal homology <DN1>
 F:2192-2351/Domains: I amino-terminal homology <DN2>
 F:2192-2345/Domains: discoidin I amino-terminal homology <DN2>
 F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1
 F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #status
 F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status experimental
 F:365-737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimental
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F:414,426/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F:759-760/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F:1667-1668/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
 F:2193-2345/Disulfide bonds: #status predicted

Query Match 7.9%; Score 105; DB 1; Length 2351;
 Best Local Similarity 23.0%; Pred. No. 3.4; Indels 76; Gaps 15;
 Matches 67; Conservative 46; Mismatches 102

QY 5 TPPIYKANKKGGKFLRLDILSPMIS-----PLGDFRHTIHIGKEGQHDVFGDISF- 58
 DB 818 TPHGSLSDLOEAKYETESDDSPGALDSNNSLSEMTHERPQLH-----HS--GDMVFT 869
 QY 59 -----LQGNVYELLPNGNEKAHLGQEPGNEFFRANSTSDSVFTETPSPYL-----KNATIS 108
 DB 870 PESGLQLRLNEKL--GTTAAWELEKLLD-----FKVSSSTNNLSTIPSDNLAAGTNDTSS 922
 QY 109 LPTIGSQALMLPLSPVTFNSKQES--FGPAKLPLRL--SCEPVMEKAQEKSSLENGTV 165

DB 923 LG-----PPMPVHYDSQDITLTLFGKSSPLTESGGPLSLSEENNDKLESLM 972
 QY 166 HQGDTWSGSSGSASOSS---QGRDSHSSSLSEQYPDMPAEDMFDPCELIKTKTSSE 222
 DB 973 NSQSSWGKNVSSITSGRLFKGRHGPALL-----TKDNALFKVISLTKTKTSNN 1025
 QY 223 SLSD---LTGSLSLQDLGLPSLDEVLN-----VMDKN 253
 DB 1026 SATNRKTHIDCP--SLIENSPPVWQNLLESDTEFKKVPFLIHDMMDKN 1074

RESULT 5
 T43225
 ptheromone receptor transcription activator - fission yeast (Schizosaccharomyces pombe
 C:Species: Schizosaccharomyces pombe
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43225; T37534
 R:Yabana, N.; Yamamoto, M.
 Mol. Cell. Biol. 16, 3420-3428, 1996
 A:Title: Schizosaccharomyces pombe map1+ encodes a MADS-box family protein required f
 A:Reference number: Z22351; MUID:96251290; PMID:8668157
 A:Accession: T43225
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-398 <YAB>
 A:Cross-references: EMBL:D78483; NID:g1754502; PIDN:BA11385.1; PID:g1754503
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsby, S.V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21721
 A:Accession: T37534
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-398 <MUR>
 A:Cross-references: EMBL:298595; PIDN:CAB1185.1; GSPDB:GN00066; SPDB:SPAC11E3.06
 A:Experimental source: strain 972h-; cosmid c11E3
 C:Genetics:
 A:Gene: map1; SPAC11E3.06
 A:Map position: 1
 A:Introns: 48/3; 105/1
 C:Keywords: transcription factor; transcription regulation

Query Match 7.5%; Score 99.5; DB 2; Length 398;
 Best Local Similarity 24.5%; Pred. No. 0.82; Indels 47; Gaps 10;
 Matches 49; Conservative 30; Mismatches 74

QY 62 NYELLPGNEKAHLGQEPGNEFFRANSTSDSVFTETPSPYLKNALSLPTIGSQALMLP 121
 DB 9 NQOI---DGEKAYTGSSOG--NSTLEDROKRONTEFTKRKAGIRKANLALTGSEVMVL- 63
 QY 122 LLSVYFNSKQESFGPAKLRLSGEPVMEKAQEKSSLENGTVHQGDTWSGSSGSASQS 181
 DB 64 ---VYSETGLVHTFSTPKLENVVKSPGOKLITES---LINATPDQNE-----SOA 108
 QY 182 SOGRDSHSSSLSEQYPDMPAEDMFH-----PPECCLIK-----GKTKSEES 223
 DB 109 SOAKQS-SAQLSDSESGP---LDHEMERISEENGPSHLENLNFSDIDNFSKTSAAEI 163
 QY 224 LSDLTGSL-----SLQDLIG 239
 DB 164 ASKLFSSVSPHTETLQFDHG 183

RESULT 6
 F90073
 hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: F90073
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
 Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of melicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2271 <KOR>
A:Cross-references: GB:BA000018; PID:913702612; PIDN:BA43752.1; GSPDB:GN00149
C:Genetics:
A:Gene: SA2447

Query Match
Best Local Similarity 7.3%; Score 97; DB 2; Length 2271;
Matches 44; Conservative 16; Mismatches 66; Indels 36; Gaps 5;

QY 85 FRANSTSDSVFTTPSPVLKNAISLPTIGSSQALMLPLSPVTFNSKQESFGPAKLPRLS 144
DB 1274 FKSESVSTSLMST-STLSNSTSLT-----SLSDSTSDSKSDSLSTM----- 1317
QY 145 CEPMEERAKQESSLLENGTVHQDTSWSSGSSASQSGRDSHSSSLSEQYPDWPAEDM 204
DB 1318 --STSDISTSKSDISTSTSLSGSTSESDSTSSSEKSDSTSMISMS----- 1366
QY 205 FDHPTPEELIKG--KTKSEESLSDLTGSLSLQLDLGPSLLD 244
DB 1367 -----QSTSGSTSTSTSTSLSDSTSTSLSTSLASMNQSGVD 1401

RESULT 7

T00530

hypothetical protein At2g19100 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T20K24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence
C:Accession: T00530; E84572
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, A.; Reference number: A84420; MUID:20083487; PMID:10617197
A:Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.
A:Reference number: Z14167
A:Accession: T00530
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1447 <ROU>
A:Cross-references: EMBL:AC002392; NID:93176701; PID:93176712
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; euss, D.; Nierman, W.C.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1447 <STO>
A:Cross-references: GB:AE002093; NID:93176712; PIDN:AAD12028.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g19100; T20K24.11
A:Map position: 2
A:Introns: 268/3; 444/1; 988/2; 1037/1; 1105/1

Query Match

Best Local Similarity 7.3%; Score 96.5; DB 2; Length 1447;
Matches 42; Conservative 23; Mismatches 74; Indels 21; Gaps 5;

QY 31 ISPLGDFRHTIHGEGQHDVFGDISFLQNYELLPGNQEKAHLGQFGHNEFRANST 90
DB 333 ITPPKTGRSPITIKTGL-----KFGVSLISNSYSVLSGDEDEVEGNKIPAGE-----ETK 384
QY 91 SDSVFTTPSP--PVLKNAISLPTIGSSQALMLPLSPVTFNSKQESFGPA--KLPRISC 145
DB 385 KDDPGTEVPYVKMTVSQDHVHVRTEVONQRLSLPRGSKTAHKVLSNSFTQSNRPFKDK 444

QY 146 EPVMEERAKQESSLLENGTVHQDTSWSSGSSASQSGR 185
DB 445 TRVKEDKAQVLSKLFN-----DWSMITNFEYNSRGR 476

RESULT 8

S26706

transcription factor SWI3 - Yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein J0495; protein YJL176c; transcription factor TYE2
C:Species: *Saccharomyces cerevisiae*
C>Date: 07-May-1993 #sequence
C:Accession: S26706; S23892; S56959
R:Peterson, C.L.; Herskowitz, I.
Cell 68, 573-583, 1992
A:Title: Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a glo
A:Reference number: S26706; MUID:92154671; PMID:1339306
A:Accession: S26706
A:Molecule type: DNA
A:Residues: 1-825 <PET>
A:Cross-references: EMBL:M84390; NID:q172799; PIDN:AAA35136.1; PID:q172800
R:Loehning, C.
Submitted to the EMBL Data Library, November 1990
A:Reference number: S23890
A:Accession: S23892
A:Molecule type: DNA
A:Residues: 1-825 <LOE>
A:Cross-references: EMBL:X56792; NID:94632; PID:94635
R:Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.
Submitted to the Protein Sequence Database, September 1995
A:Reference number: S56937
A:Accession: S56959
A:Molecule type: DNA
A:Residues: 1-825 <OBE>
A:Cross-references: EMBL:Z49451; NID:q1008372; PID:q1008373; MIPS:YJL176c
C:Genetics:
A:Gene: SGD:SWI3; TYE2
A:Cross-references: SGD:S0003712; MIPS:YJL176c
A:Map position: 10L
C:Function:
A:Description: transcription regulation
C:Keywords: nucleus; transcription regulation

Query Match

Best Local Similarity 7.2%; Score 96; DB 2; Length 825;
Matches 50; Conservative 31; Mismatches 75; Indels 42; Gaps 9;

QY 69 NOEKAHLGQFGHNEFRANSTSDSVFTTPSPVLKNAISLPTIGSSQALMLPLSPV-T 127
DB 88 SQETPSTESRAQVYFGQDNEDSDNLFGETESSVNNNEANTPSIP-----TNPVDN 138
QY 128 FNSK-----QESFGPAKLPRLSCEPVMEERAKQESSLLENGTVHQDTSWSSGSA 178
DB 139 ENNKPAIKEDSTIQDSNGDK-----NMEDVKIQKEEP-----ENNTVIEG-----V 181
QY 179 SQSSQGRDSHSSSLSEQYPDWPAEDMHP--TPCELLKGTKSEESLSDLTGSLSLQL 236
DB 182 KEESQ-PDENTKEMDEVED--DEDDQPMISPDNSIFGDTKSESKQLGNTSSVANTPS 237
QY 237 DLGSLDEVLNVMDRKNK 254
DB 238 EIPDAHKAQEDITEKTE 255

RESULT 9

T42013

frequency clock protein - *Creopus spinulosus*
C:Species: *Creopus spinulosus*
C>Date: 03-Dec-1999 #sequence
C:Accession: T42013
R:Lewis, M.T.; Morgan, L.W.; Feldman, J.F.
Mol. Gen. Genet. 253, 401-414, 1997
A:Title: Analysis of frequency (frq) clock protein homologs: evidence for a helix-turn
A:Reference number: Z23024; MUID:97188515; PMID:9037100

A:Accession: T42013
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1015 <LEW>
A:Cross-references: EMBL:U25850; NID:9852501; PID:9852502; PIDN:AAA68072.1
A:Gene: frq
A:Introns: 100/1

Query Match 7.2%; Score 95.5; DB 2; Length 1015;
Best local similarity 19.6%; Pred. No. 6.8; Indels 97; Gaps 10;
Matches 55; Conservative 40; Mismatches 97

QY 26 LSPDMISPLGDFRHTHIGKEGQ-----HDVFGDISFLQNGYELPQNGE 581
DB 522 VTPDFVRNAVYDLAKFQLSPGKRIRRGCTDGTKESSSSGDLSPETDDTENTKG 131
QY 72 KAHLGQFPGHNEFRANSTDSVFTETPSPVLKNAISLPTIGSQALMLPLSPVTENSK 614
DB 582 DNHRKOKTGH-----STGDSG-----SSGNNLPKFGGPV-----SAS 614
QY 132 QESFGPAKLPRLSCEPYMEKAQEKSSILENGTV-----HOGDTWSSGSSASQSSQ 183
DB 615 SESFHYKPL-----FLHQOSPNEQSSMEDGTLSSFGPIESNADSRWGQSGSASNR 667
QY 184 GR-----DSHSSLSLSEQPDWPAEDMDHPTPCELIKG 216
DB 668 KRRRDGAIIYSGAPFCTDLSDPGDTPATYMLSSERKRPD-AQGQFARPLPFRSSG 725
QY 217 KTKSESLSDL-----TGSLLSLQ-LDLGSLDLDEVLNVMKN 253
DB 726 SSITRRPLSDAHLNLSIPKQLLNQVSIPELVTDGDES 766

RESULT 10

T22981
hypothetical protein F59B10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22981

R:loyd, C.
Submitted to the EMBL Data Library, March 1995
A:Reference number: Z19646
A:Accession: T22981
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-482 <WIL>
A:Cross-references: EMBL:Z48716; PIDN:CAA88601.1; GSPDB:GN00020; CESP:F59B10.2
A:Experimental source: clone F59B10
C:Genetics:
A:Gene: CESP:F59B10.2
A:Map position: 2
A:Introns: 40/2; 74/1; 118/3; 415/1

Query Match 7.2%; Score 95; DB 2; Length 482;
Best local similarity 23.6%; Pred. No. 2.6; Indels 99; Gaps 12;
Matches 65; Conservative 33; Mismatches 99

QY 1 MPAPTPIYKANNKKGKFF-----KLDELSPDMISPLGDFRHTHIGKEG-QHDFEG 54
DB 68 LPLGTP-----GOQVGRSLKIDIRELRNSIDSLIKPPELKMRSKAAGEESKRAHFE 121
QY 55 DISFLQNGYELLPNGOEKAHIGQFPGHNEFFRA-----NSTSDS-----VFTET-PSPV 102
DB 122 SIQEEKISIEDLP-----TAQLPSKSPFKKAISGSSSSDSITRDEVEEVLPSPP 174
QY 103 LKNAISLPTIGSQALMLPLSPVTENSKQESFGPAKLPRLSCEPYMEKAQEK-----156
DB 175 RKPAPA-----RTAPIVEKRIEK--PA-----VKEQARKKKEKPTPT 210
QY 157 -----SSLNGLTVHOGDTWSSGSSASQSGGRDSSSLSEQPDWPAEDMDHPTPT 210
DB 211 PTSSFFSSSSDSSSTSSSTSSSSSSASSESSESSESQVSSSKTSTKASSSKAYGSD 270

QY 211 CELIKGKTKSEESLSDLTGSLSLSLQDLGSLDE 245
DB 271 FESEKSSSSASTISKVT-----PKKLDK 294

RESULT 11

T45793
hypothetical protein F26013.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45793
R:Delaney, M.; Berger, C.; Cooke, R.; Grellet, F.; Laurie, M.; Mewes, H.W.; Lemcke, K
Submitted to the Protein Sequence Database, December 1999
A:Reference number: 223013
A:Accession: T45793

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-764
A:Cross-references: EMBL:AL133452
A:Experimental source: cultivar Columbia; BAC clone F26013
C:Genetics:
A:Map position: 3
A:Introns: 76/3; 138/3; 169/3; 204/1; 266/1; 621/3
A:Note: F26013.260

Query Match 7.1%; Score 94; DB 2; Length 764;
Best local similarity 22.7%; Pred. No. 6.1; Indels 92; Gaps 17;
Matches 66; Conservative 35; Mismatches 92

QY 9 LKAANN-----KKGKFKLRD-----ILSPDMISPLGDF-----RHTIHI 44
DB 253 LKETNNLGRSVSKGNEYRIRSAFTYGARKLGQLFLOSDAISELSKRFKFSMLLRH-----308
QY 45 GKEGQDVFGDISFLQNGYELLPNGN-----QEKALGQFPGHNEFRANSTDSVFT-----96
DB 309 GSGQRPDVHDAIPFLRYNRYNAILPASNHFE-----GOVNESE-----SSSSGATGNG 359
QY 97 -ETPSPVLKNAISLPT-----IGSQALMLPLSPVTENSKQESFGPAKLPRLSC-----145
DB 360 RHDQEDSLDAGVISPTTGPDLGSGPGETVPVSSEERFSGDAKLATRIOKLEISDDAM 419
QY 146 -EPVMEKAQEK-----SSLNGLTVHOGDTWSSGSSASQSGGRDSSHS 189
DB 420 KSPCLSDKESDSPLNKXHSFNQMRNGEVLNGVKGQENSWHTG-----SRVYKDIHI 474
QY 190 SSLSEQ-----YPD-----WPAEDMDHPTPCELIKGKTKSEESLSDLTG 229
DB 475 NENENHGVYEDLPFASAVPWPQEDMHLHYSG-HCVSG-----TPNMLSDLSG 521

RESULT 12

A84616
hypothetical protein At2g22720 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84616
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-598 <STO>
A:Cross-references: GB:AE002093; NID:94314369; PIDN:AAD15580.1; GSPDB:GN00139
A:Map position: 2

Query Match 7.1%; Score 93.5; DB 2; Length 598;

Best Local Similarity 21.7%; Pred. No. 4.8;
Matches 51; Conservative 26; Mismatches 107; Indels 51; Gaps 7;

QY 55 DISFL-QGNYEL-LPGNOEKAHLGQFPNGHNEFFRANS-----
Db 166 DYSLFSDDELFPVKESLSRSGSFPNSAYHFEHEDNLYRFADYQEARSAQLSSRPKOS 89
QY 90 -----TSDSVFTETPPSPVLKNAISLPTIGSQALMLPLSPVTFNSKQESFGPAKL--- 140
Db 226 SGINGRTAHPHREKRPVSANGHSRPSGSSQ-----MNSRPSGSSGSKMNS 274
QY 141 -PRLSCEPVMEEKAEKSSLENGTVHQDTSWSSGASQSSQGRDSSSLSEQYPDW 199
Db 275 RPATSGSQMPNRPASSGQMSRAV-SGSGRPASSGQMSRPNQSRPASAGSQMQQR 333
QY 200 PAEDMFHPTPCELIKGKTKSEESLSDLTGSLSLQDLGSLIDEVLNMDKNK 254
Db 334 PASSGQRPAS-----SGSGRPASSGQMSRPNQSRPASAGSQMQQR 384

RESULT 13

WMHUE2

HIV-EP2 enhancer-binding protein - human
N:Alternate names: finger protein, 275K; human immunodeficiency virus enhancer-binding
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence_revision 07-Jul-1995 #text_change 21-Jul-2000
R:Van't Veer, L.J.; Lutz, P.M.; Isselbacher, K.J.; Bernards, R.
A:Title: Acad. Sci. U.S.A. 89, 8971-8975, 1992
A:Reference number: S26661; MUID:93028387; PMID:1409593
A:Accession: S26661
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2500 <VAN>

R:Nomura, N.; Zhao, M.J.; Nagase, T.; Maekawa, T.; Ishizaki, R.; Tabata, S.; Ishii, S.
J. Biol. Chem. 266, 8590-8594, 1991
A:Title: HIV-EP2, a new member of the gene family encoding the human immunodeficiency vi
A:Reference number: A39829; MUID:91217105; PMID:2022670
A:Accession: A39829
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 668-2144, 'R', 2146-2500 <NOM>
A:Cross-references: GB:M60119; NID:92661140; PIDN:AAB88218.1; PID:9182120
R:Rustgi, A.K.; Van't Veer, L.J.; Bernards, R.
Proc. Natl. Acad. Sci. U.S.A. 87, 8707-8710, 1990
A:Title: Two genes encode factors with NF-kappaB- and H2TF1-like DNA-binding properties
A:Reference number: A38253; MUID:91062349; PMID:2247438
A:Accession: A38253
A:Molecule type: mRNA
A:Residues: 1851-1990 <RUS>
A:Cross-references: GB:M61744; GB:M33920; NID:9187404; PIDN:AAA36202.1; PID:9187405
C:Genetics:
A:Gene: GDB:HIVEP2
A:Cross-references: GDB:129086; OMIM:143054
A:Map position: 6q23-6q24
C:Superfamily: HIV-EP2 enhancer-binding protein
C:Keywords: DNA binding; duplication; metal binding; transcription regulation; zinc fing

F:242-298/Region: DNA binding; duplication; metal binding; transcription regulation; zinc fing
F:245-265/Region: zinc finger CCH motif
F:273-295/Region: zinc finger CCH motif
F:991-997/Region: nuclear location signal
F:1004-1036/Region: serine-rich
F:1852-1908/Region: DNA binding #status predicted
F:1855-1875/Region: zinc finger CCH motif
F:1883-1905/Region: zinc finger CCH motif
F:1953-1977/Region: acidic

Query Match
Best Local Similarity 26.7%; Score 93.5; DB 1; Length 2500;
Matches 50; Conservative 24; Mismatches 66; Indels 47; Gaps 7;

QY 89 STSDSVFTETPPSPVLKNAISLPTIGSQALMLPLSPVTFNSKQESFGPAKLPRISCEP 147
Db 1481 STSDSV-----ATLGSKRMLSPASSLELFEMETKQK--RVKEEMYGQI 1523
QY 148 VMEEKAE-----KSLLENGTVHQDTSWSSGASQSSQGRDSSSLSEQYPDW 199
Db 1524 VEELSAVELTNSDIKDLSPQKPOLVRQGCASEPRKDLQSSSSFSLSPPSSQDYPV 1583
QY 191 SLSEQYPDWPAEDMF-----HPTPCELIKGKTKSEESLSDLTGSLSLQDL 238
Db 1584 SPSSREPPFPKEMLSGRAPLPQKSSGSPSEKSSDELIDETAASDMSMSPSSSLPA 1643
QY 239 GPSTLDE 245
Db 1644 GDQLE 1650

RESULT 14

T17202

DNA-directed DNA polymerase (EC 2.7.7.7) zeta chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Van Sloun, P.P.H.; Romeijn, R.J.; Beken, J.C.J.
Mutat. Res. 433, 109-116, 1999
A:Title: Molecular cloning, expression and chromosomal localisation of the mouse Rev3
A:Reference number: Z18720; MUID:99202265; PMID:10102037
A:Accession: T17202
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3122 <VAN>

A:Cross-references: EMBL:AF083464; NID:94079830; PID:94079831; PIDN:AAC98785.1
A:Experimental source: strain 129/Ola; testis
C:Genetics:
A:Map position: 10
C:Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 24.5%; Score 93.5; DB 2; Length 3122;
Matches 58; Conservative 30; Mismatches 104; Indels 45; Gaps 12;

QY 9 LKAANKKKKKFKRLDILSPDM-----ISPLGDFRTHHGKEGOHDFGDISFLQCN 63
Db 718 LSSGNEKGNSTSLSGVFPSSLTENCILLPSSGEMRSMH-SLESTDESGLNKLKIRY 775
QY 64 ELLPGNO-EKAHLGQFPNGHNEFFRANSTSDSVFTETPPSPVLKNAISLPTIGSQALMLPL 122
Db 776 EEFQEHKMEKPSLSQAAHYMF-----PSVLSNCLTRP-----QK 812
QY 123 LSPVTFNSKQESFGPAKLPRISCEPVMEEKAEKSSLENGTVHQDTSWSSGASQSS 182
Db 813 LSPVTF-QLSGNKPRLKLNKKKLGIGETSTKST-ETGATKDSCTHNDLYTGASEKE 869
QY 183 QGRDSSSSLS-----EQYPDWPAEDMFHPTPCELIKGKTKSEESLSDLTGSLSLQ 235
Db 870 NGLSSDSAKATGTFENKP--PTEHFTD-----CHFGDGSLEAEQSFGLYGNKTYLR 919

RESULT 15

T51880

hypothetical protein DKFZp762G2015.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
R:Ansorge, W.; Winkler, U.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, July 2000
A:Accession: T51880
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-606 <AAA>
A:Cross-references: EMBL:AL390166
A:Experimental source: adult melanoma (Mewo cell line); clone DKFZp762G2015

Query Match
Best Local Similarity 24.5%; Score 93.5; DB 2; Length 3122;
Matches 58; Conservative 30; Mismatches 104; Indels 45; Gaps 12;

Wed Apr 23 13:51:14 2003

us-09-806-276a-2.rpr

C:Genetics:
A>Note: DKFZp762G2015.1

7.08; Score 93; DB 2; length 606;
Query Match 21.7%; Pred. No. 5.4; Indels 76; Gaps 11;
Best Local Similarity 38; Mismatches 106; Indels 76; Gaps 11;
Matches 61; Conservative 38; Mismatches 106; Indels 76; Gaps 11;

```
44 IGKGGHDFVFGDIS---FLQGNVELLPNGOEKA-----HLG-----Q 77
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 VMQGGKGVAYDLGRVERIMSETYMLIKQVDEEALQAVKFCQVHLGLPPDRPRDPT 367
QY 78 FPGHNEFFRANSTSDSVFTETPSPVLKNAISLPT-----IGSQALML--PLSPV 126
Db 368 TPKHPKDSRENEFFVTVPPTAPDPVPADSVQRPDAHTKPPALAAATTTTCPPSASAS 427
QY 127 TFNSKQESFGPAKLPR-----LSCEPVMEKAQEKSSLLENGTVHQDTS 171
Db 428 TLDQSKDS-GPPRPHRPEATPSMASLCPBEGKLRPERRRDEQAQEAAS--ETQPLSSPPTA 484
QY 172 WSSSGSASQSSQGRDSH-----SSSLSPQYPD-----WPAEDMFDHPTPCEL 213
Db 485 ASSKAPSSGSAQPEGHGPKPEPSRAKSRPLPMKLVIPSAATKFPPEITVTPPTPTLL 544
QY 214 IKGKTKSEESLDTGLSLSLQDLG---PSLDEVLNMD 251
Db 545 SPKGSISETKOKLKSAILSAQSAANVRKESLCPALEVLE 585
```

Search completed: April 23, 2003, 11:03:23
Job time : 20.0943 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:54:35 ; Search time 15.0943 Seconds
(without alignments)
697.946 Million cell updates/sec

Title: US-09-806-276A-2
Perfect score: 1326
Sequence: 1 MPAKTPYILKANNKKKKF.....QLDLGPSLLDEVINVMKDK 254

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222.5	16.8	356	1 BOR4_HUMAN	Q9h3q1 homo sapien
2	221	16.7	349	1 BOR4_MOUSE	Q9jmg6 mus musculu
3	195	14.7	391	1 M5E5_HUMAN	Q00587 homo sapien
4	109.5	8.3	1475	1 N153_HUMAN	P49790 homo sapien
5	105	7.9	2351	1 FA8_HUMAN	P00451 homo sapien
6	104	7.8	1805	1 RW1_HUMAN	Q92545 homo sapien
7	99.5	7.5	398	1 MAP1_SCHPO	P78926 schizosacch
8	96	7.2	825	1 SWI3_YEAST	P32591 saccharomyc
9	95.5	7.2	1015	1 FRO_CRESP	Q00586 creopus spi
10	95	7.2	482	1 YSR2_CAEEL	Q00950 caenorhabdi
11	93.5	7.1	1833	1 ZEP2_HUMAN	P31629 homo sapien
12	93.5	7.1	3122	1 DPO2_MOUSE	Q61493 mus musculu
13	92.5	7.0	1337	1 YDM5_SCHPO	P87136 schizosacch
14	92	6.9	667	1 ITF2_HUMAN	P15884 homo sapien
15	90.5	6.8	899	1 YABD_SCHPO	Q09778 schizosacch
16	90.5	6.8	1128	1 BEM3_YEAST	P32873 saccharomyc
17	90	6.8	633	1 ACES_ELEEL	O42275 electrophor
18	89.5	6.7	510	1 YDID_SCHPO	O13692 schizosacch
19	89.5	6.7	1612	1 DNMI_PARLI	Q27746 paracentrot
20	89	6.7	289	1 MYOD_XENLA	P13904 xenopus lae
21	89	6.7	1744	1 TENS_CHICK	Q04205 gallus gall
22	88.5	6.7	642	1 ITF2_CANFA	P15881 canis famli
23	88.5	6.7	741	1 GTSE_MOUSE	O8r080 mus musculu
24	88	6.6	499	1 GAG_HV1J3	P12494 human immun
25	86	6.5	478	1 JTM3_YEAST	P47018 saccharomyc
26	86	6.5	617	1 SPH2_MOUSE	O9j1a7 mus musculu
27	86	6.5	1142	1 MGCI_HUMAN	O60732 homo sapien
28	86	6.5	1347	1 UTY_HUMAN	O14607 homo sapien
29	85.5	6.4	1189	1 YJH6_YEAST	P47035 saccharomyc
30	85.5	6.4	1237	1 E75A_DROME	P17671 drosophila
31	85.5	6.4	1394	1 E75B_DROME	P17672 drosophila
32	85.5	6.4	3256	1 K167_HUMAN	P46013 homo sapien
33	85	6.4	660	1 CCB2_HUMAN	Q08289 homo sapien

34	85	6.4	2805	1 MAPA_HUMAN	P78559 homo sapien
35	84.5	6.4	611	1 IF4B_HUMAN	P23588 homo sapien
36	84.5	6.4	995	1 YI09_YEAST	P40442 saccharomyc
37	84	6.3	632	1 CCB2_RABIT	P54288 oryctolagus
38	84	6.3	1004	1 SAL2_MOUSE	Q9qx96 mus musculu
39	84	6.3	2148	1 VITI_AEDAE	Q16927 aedes aegypt
40	84	6.3	2774	1 MAPA_RAT	P34926 ratius norv
41	83.5	6.3	247	1 TRY2_HUMAN	P07478 homo sapien
42	83.5	6.3	356	1 SMRI_PODAN	Q08142 podospira a
43	83.5	6.3	470	1 PLSB_CUCSA	Q36639 cucumis sat
44	83.5	6.3	1268	1 PGCN_MOUSE	P55066 mus musculu
45	83	6.3	278	1 YD33_SCHPO	Q10267 schizosacch

ALIGNMENTS

RESULT 1
BOR4_HUMAN STANDARD; PRT; 356 AA.

AC Q9H3Q1; O95828; Q96FT3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Binder of Rho GTPase 4 (Cdc42 effector protein 4).
GN BOR4 OR CEP4.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

[1]
SEQUENCE FROM N.A.
TISSUE=Ileal mucosa;
MEDLINE=21036164; PubMed=11185749;
Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
"Sequence analysis, gene expression, and chromosomal assignment of
mouse Borg4 gene and its human orthologue."
J. Hum. Genet. 45:374-377(2000).

[2]
SEQUENCE FROM N.A., AND CHARACTERIZATION.

TISSUE=Eye;
MEDLINE=21125609; PubMed=11035016;
Hirsch D.S., Pirone D.M., Burelo P.D.;
"A new family of cdc42 effector proteins, CEPs, function in fibroblast
and epithelial cell shape changes."
J. Biol. Chem. 276:875-883(2001).

[3]
SEQUENCE FROM N.A.

TISSUE=Brain;
Strausberg R.;
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Probably involved in the organization of the actin
cytoskeleton. May act downstream of CDC42 to induce actin
filament assembly leading to cell shape changes. Induces
pseudopodia formation, when overexpressed in fibroblasts.
-1- SUBUNIT: Interacts with CDC42 and TC10, in a GTP-dependent manner
(By similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated.
-1- TISSUE SPECIFICITY: Not detected in any of the adult tissues
tested. May be expressed only in fetal or embryonic tissues.
-1- SIMILARITY: BELONGS TO THE BORC/CEP FAMILY.
-1- SIMILARITY: CONTAINS 1 CRIB DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AB042237; BAB17272.1; -
EMBL; AF099664; AAD16299.1; -
DR

EMBL; BC010451; AAH10451.1; -
DR MIM; 605468; -
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
DR PROSITE; PS50108; CRIB; 1.
KW Cytoskeleton.
FT DOMAIN 27 41 CRIB.
FT DOMAIN 70 74 POLY-SER.
FT CONFLICT 2 2 P -> L (IN REF. 3).
FT CONFLICT 33 33 D -> T (IN REF. 2).
FT CONFLICT 288 288 A -> T (IN REF. 2).
FT CONFLICT 296 296 A -> T (IN REF. 2).
FT CONFLICT 339 339 P -> P (IN REF. 2).
FT CONFLICT 355 355 MISSING (IN REF. 2).
SQ SEQUENCE 356 AA; 37979 MW; 2CF677C60C6EF1B5 CRC64;

Query Match
Best Local Similarity 16.8%; Score 222.5; DB 1; Length 356;
Matches 77; Conservative 37; Mismatches 91; Indels 67; Gaps 11;

QY 6 PLYKANKKGGKFKRLDILSPDMISPLGDFRHTIHGKEGQHDVFGDISFLQGNTEL 65
DB 2 PILKQLVSSSVHSKRSRADLTAEWISAPLGDFRHTMHVGRAG--DAFGDTSLNSK--- 56
QY 66 LFGNQEKALHGFPGHNEF-----FRANSTSDSV-----FTETPSVLYKN 105
DB 57 -AGEPDGESLDEPDSSSSSKRSLSRKFRGSKRSQSVTRGERQRMGLSLRDSALFVK 115
QY 106 AISLPTIGSQALMLPLSPVTENSKQESFGPAKLPR-LSCPEVMEKAQEKSSLENGT 164
DB 116 AMSLPQNEKA-----AEKGTSLPKSLSSSPV--KKANDGEG----- 152
QY 165 VHOGDTSMGSSGASQSSQGRDSSSS--LSEQYPDWPAEDMFDHPTPCELIKGKTSEE 222
DB 153 ---GDEEAGTEAVPRRNGAGPSPDPLDEQ---AFGDLTDLPVVPRKATYGLKHA 204
QY 223 SLSDLTGLSLQLDLGPSLDEVLVNMDKNK 254
DB 205 -----SISMFHIDLGPMLGDLVLSIMDKKE 229

RESULT 2
BOR4_MOUSE
ID BOR4_MOUSE STANDARD; PRT; 349 AA.
AC Q9JUM96; Q9JUM96; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN Blinder of Rho GTPase 4 (Cdc42 effector protein 4).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=ileal mucosa;
RX MEDLINE=21036164; PubMed=11185749;
RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
RT "Sequence analysis, gene expression, and chromosomal assignment of
mouse Borg4 gene and its human orthologue.";
RL J. Hum. Genet. 45:374-377(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-143 FROM N.A., AND INTERACTION WITH TC10 AND CDC42.
RX MEDLINE=99421943; PubMed=10490598;
RA Jobery G., Perlungher R.R., Macara I.G.;
RT "The Borgs, a new family of Cdc42 and TC10 GTPase-interacting
proteins.";
RL Mol. Cell. Biol. 19:6585-6597(1999).

CC -1- FUNCTION: Probably involved in the organization of the actin
cytoskeleton. May act downstream of CDC42 to induce actin
filament assembly leading to cell shape changes. Induces
pseudopodia formation, when overexpressed in fibroblasts.
CC -1- SUBUNIT: Interacts with CDC42 and TC10, in a GTP-dependent manner.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated (By
similarity).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 CRIB DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AB035088; BAA95932.1; -
DR EMBL; BC003857; AAH03857.1; -
DR EMBL; AF165114; AAD47822.1; -
DR MGD; MGI:1929760; Borg4.
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
DR PROSITE; PS50108; CRIB; 1.
KW Cytoskeleton.
FT DOMAIN 27 41 CRIB.
FT CONFLICT 143 143 K -> R (IN REF. 3).
SQ SEQUENCE 349 AA; 37869 MW; 7C44125A7083E16B CRC64;

Query Match
Best Local Similarity 16.7%; Score 221; DB 1; Length 349;
Matches 76; Conservative 42; Mismatches 83; Indels 66; Gaps 10;

QY 6 PLYKANKKGGKFKRLDILSPDMISPLGDFRHTIHGKEGQHDVFGDISFL-----Q 60
DB 2 PILKQLVSSSVHSKRSRADLTAEWISAPLGDFRHTMHVGRAG--DAFGDTSLNSKARE 59
QY 61 GNYELLPGNQEKALHGFPGHNEFFRANSTSDSV-----FTETPSVLYKN 108
DB 60 ADDESL---DEQASASKLSLSRFRGSKRSQSVTRGERQRMGLSLRDSALFVK 116
QY 109 LPTIGSQALMLPLSPVTENSKQESFGPAKLPR-LSCPEVMEKAQEKSSLENGTVHQ 167
DB 117 LPQNEKA-----AEKDS--SKLPKSLSSSPVKKADARD----- 149
QY 168 GDTSMGSSGASQSSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSD 227
DB 150 ---GDEKSPHRNGATGPHSPDPLDEQ---AFGDLMDLPIMPKYVSYGLKHA 195
QY 228 TGLSLQLDLGPSLDEVLVNMDKNK 254
DB 196 --SISMFHIDLGPMLGDLVLSIMDKDQ 220

RESULT 3
MSE5_HUMAN
ID MSE5_HUMAN STANDARD; PRT; 391 AA.
AC O00587; Q96GNT; 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN Serum protein MSE5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=92332498; PubMed=1629197;
RA Bahou W.F., Campbell A.D., Wicha M.S.;

DR PIR; A429/3; A425/5.

entitles requires a license agreement (See <http://www.isb.org>)

or send an email to license@sib.ch.

CC -----
DR EMBL; Z25535; CAA80982.1; -
DR Genew; HGNC:8062; NUP153.
DR MIM; 603948; -
DR InterPro; IPR001876; znf_RangDP.
DR Pfam; PF00641; zf_RanBP; 4.
DR SMART; SM00547; znf_RB2; 4.
DR PROSITE; PS01358; zf_RANBP2_1; 4.
DR PROSITE; PS50199; zf_RANBP2_2; 4.
KW Nuclear protein; Transport; Repeat; Zinc-finger; DNA-binding.
FT DOMAIN 4 14
FT DOMAIN 4 14
FT ZN_FING 443 447 POLY-GLY.
FT ZN_FING 657 687 RANBP2-TYPE 1.
FT ZN_FING 722 751 RANBP2-TYPE 2.
FT ZN_FING 793 822 RANBP2-TYPE 3.
FT ZN_FING 851 880 RANBP2-TYPE 4.
SQ SEQUENCE 1475 AA; 153889 MW; 3CB415A6909DF80E CRC64;

Query Match 8.3%; Score 109.5; DB 1; Length 1475;
Best Local Similarity 26.0%; Pred. No. 0.48;
Matches 57; Conservative 31; Mismatches 86; Indels 45; Gaps 10;

QY 32 SPPLEPRHTI-----HIGEGQHD--VFGDISFLQNYELLPGNQEKALHGQFPQHN 82
Db 953 SKPIGDKFKGVSSSEKPEEVKKDSKNDKFKGLSSGLSNPVSILTPFGVSNLGOEKEKE 1012
QY 83 EFRANSTSDS---VETETSPV-----LKNATSLPTIGSSQALMLPLSPVTFNSK 131
Db 1013 ELPKSSSAGFSFGTGVINSTPAPANTIVTSENKSSFNLTETKASAVAPFTCKTSEAKK 1072
QY 132 QE-----SFG---PAKLPLRSCPEV--MEKAQE--KSSLNGTGVHOGDT----- 170
Db 1073 EEMPATKGGTSGNVEPASPASVFLGRTEKQEQEPTVSTSLVEGKKADNEPRCQPV 1132
QY 171 -SWGSSGSASQSSQGRDSSHSSSL---SEQYDWPAPEDMF 205
Db 1133 FSFGNSEQTKDENSSTKSTFSMTKPSKESEQPAKATF 1171

RESULT 5

FA8_HUMAN
ID FA8_HUMAN STANDARD; PRT; 2351 AA.
AC P00451;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component)
DE (Antihemophilic factor) (AHF).
GN F8 OR F8C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Maslarsz F.R., Merryweather J.P., Najarian R.,
RA Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA Nordfang O., Ezban M.;
RT "Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA."
RL DNA 4:333-349(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wilson K.L.,
RA Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RT "Expression of active human factor VIII from recombinant DNA clones.";

RL Nature 312:330-337(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Soltzman L.A., Buckner J.L.,
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA Hewick R.M.;
RT "Molecular cloning of a cDNA encoding human antihemophilic factor."
RT Nature 312:342-347(1984).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93265012; PubMed=1303178;
RA Gitschier J., Wood W.I.;
RT "Sequence of the exon-containing regions of the human factor VIII
RT gene."
RT Hum. Mol. Genet. 1:199-200(1992).
RN [5]
RP SEQUENCE OF 2064-2070 FROM N.A.
RA de Water N.S., Williams R., Browett P.J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SUFFATION OF TYR-1699.
RX MEDLINE=91093266; PubMed=1898735;
RA Leyte A., van Schijndel H.B., Niehrs C., Rutner W.B., Verbeet M.P.,
RA Mertens K., van Mourik J.A.;
RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
RT essential for the interaction of factor VIII with von Willebrand
RT J. Biol. Chem. 266:740-746(1991).
RN [7]
RP SUFFATION.
RX MEDLINE=92207952; PubMed=1554716;
RA Pittman D.D., Wang J.H., Kaufman R.J.;
RT "Identification and functional importance of tyrosine sulfate
RT residues within recombinant factor VIII."
RL Biochemistry 31:3315-3325(1992).
RN [8]
RP STRUCTURE BY NMR OF 2322-2343.
RX MEDLINE=95200924; PubMed=7893714;
RA Gilbert G.E., Baleja J.D.;
RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT amphipathic structure as determined by NMR spectroscopy."
RL Biochemistry 34:3022-3031(1995).
RN [9]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE=91221499; PubMed=1902642;
RA Gitschier J.;
RT "The molecular basis of hemophilia A."
RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
RN [10]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE=89088506; PubMed=2491949;
RA White G.C. II, Shoemaker C.B.;
RT "Factor VIII gene and hemophilia A."
RL Blood 73:1-12(1989).
RN [11]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE=95245332; PubMed=7728145;
RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
RT "Molecular etiology of factor VIII deficiency in hemophilia A."
RL Hum. Mutat. 5:1-22(1995).
RN [12]
RP VARIANT GLN-2326.
RX MEDLINE=86235434; PubMed=3012775;
RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
RT "Identification of a missense mutation in the factor VIII gene of a
RT mild hemophilic."
RL Science 232:1415-1416(1986).
RN [13]
RP VARIANT PRO-2135.
RX MEDLINE=88096539; PubMed=3122181;
RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;

Wed Apr 23 13:51:15 2003

"A novel missense mutation in the factor VIII gene identified by analysis of amplified hemophilia DNA sequences.";
 Nucleic Acids Res. 15:9797-9805(1987).
 [14]
 VARIANT GLN-2228. PubMed=2833855;
 MEDLINE=88191889; Antonarakis S.E., Bell W., Griffin A.M., Youssoufian H., Antonarakis S.E., estimate of the Kazazian H.H.;
 "Nonsense and missense mutations in hemophilia A: estimate of the relative mutation rate at CG dinucleotides.";
 Am. J. Hum. Genet. 42:718-725(1988).
 [15]
 VARIANT GLY-291. PubMed=2835904;
 MEDLINE=88220354; Platokoukis H., Kazazian H.H. Jr., Youssoufian H., Wong C., Aronis S., platokoukis H., Kazazian H.H. Jr., Antonarakis S.E.;
 "Moderately severe hemophilia A resulting from Glu-->Gly substitution in exon 7 of the factor VIII gene.";
 Am. J. Hum. Genet. 42:867-871(1988).
 [16]
 VARIANT CYS-1708. PubMed=2499363;
 MEDLINE=89274393; Tuddenham E.G.;
 O'Brien D.P., Tuddenham E.G.;
 "Purification and characterization of factor VIII L, 689-Cys: a nonfunctional cofactor occurring in a patient with severe hemophilia A.";
 Blood 73:2117-2122(1989).
 [17]
 VARIANT CYS-391. PubMed=2506948;
 MEDLINE=90001543; Fukui H., Fulcher C.A.;
 Shima M., Ware J., Yoshioaka A., Fukui H., Fulcher C.A.;
 "An arginine to cysteine amino acid substitution at a critical thrombin cleavage site in a dysfunctional factor VIII molecule.";
 Blood 74:1612-1617(1989).
 [18]
 VARIANT LEU-189. PubMed=2510835;
 MEDLINE=90057680; Todd D.;
 Chan V., Chan T.K., Tong T.M., Todd D.;
 "A novel missense mutation in exon 4 of the factor VIII:C gene resulting in moderately severe hemophilia A.";
 Blood 74:2688-2691(1989).
 [19]
 VARIANT LEU-2326. PubMed=2495245;
 MEDLINE=89197216; Kazazian H.H. Jr., Antonarakis S.E.;
 Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 "Wild hemophilia A resulting from Arg-to-Leu substitution in exon 26 of the factor VIII gene.";
 Hum. Genet. 81:335-338(1989).
 [20]
 VARIANT HIS-391. PubMed=2498882;
 MEDLINE=89264602; Antonarakis S.E., Kazazian H.H. Jr., Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr., Fujimaki M., Hoyer L.W.;
 "Direct characterization of factor VIII in plasma: detection of a mutation altering a thrombin cleavage site (arginine-372-->histidine).";
 Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 [21]
 VARIANT CYS-1708. PubMed=2104766;
 MEDLINE=90105723; Kazazian H.H. Jr., Antonarakis S.E., Higuchi M., Antonarakis S.E., Hoyer L.W.;
 Arai M., Higuchi M., Antonarakis S.E., Hoyer L.W.;
 "Characterization of a thrombin cleavage site mutation (Arg 1689 to Cys) in the factor VIII gene of two unrelated patients with cross-reacting material-positive hemophilia A.";
 Blood 75:384-389(1990).
 [22]
 VARIANTS GLN-2228 AND LEU-2326. PubMed=2105106;
 MEDLINE=90123183; Ristaldi M.S., Restagno G., Casula L., Murru S., Pecorara M., Ristaldi M.S., Carbonara A.;
 Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 "Recurrent mutations and three novel rearrangements in the factor VIII gene of hemophilia A patients of Italian descent.";

Blood 75:662-670(1990).
 [23]
 VARIANT CYS-391. PubMed=1973901;
 MEDLINE=90329422; McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the internal heavy chain thrombin cleavage site.";
 Br. J. Haematol. 75:73-77(1990).
 [24]
 VARIANTS PHE-1699 AND CYS-1708. PubMed=2105906;
 MEDLINE=90152691; Olek K., Aronis S., Kasper C.K., Higuchi M., Wong C., Kochan L., Olek K., Aronis S., Kasper C.K., Kazazian H.H., Antonarakis S.E.;
 "Characterization of mutations in the factor VIII gene by direct sequencing of amplified genomic DNA.";
 Genomics 6:65-71(1990).
 [25]
 VARIANTS CYS-1728 AND ASP-1941. PubMed=2106480;
 MEDLINE=90169988; Kasper C.K., Antonarakis S.E., Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E., Kazazian H.H.;
 "Use of denaturing gradient gel electrophoresis to detect point mutations in the factor VIII gene.";

Query Match 7.9%; Score 105; DB 1; Length 2351;
 Best Local Similarity 23.0%; Pred. No. 2.1; 102; Indels 76; Gaps 15;
 Matches 67; Conservative 46; Mismatches 58

QY 5 TPVYKANNKKKKFKLRDILSPDMIS-----PPLGDFRHTTHIGKEGQHDVFGDISF- 58
 818 TPHGLSLDQEAKEYETFSDDPSGAIDSNLSSEKTHRPQLH-----HS--GDMVFT 869
 QY 59 -----LQGNVELLPNGOEKAHLGQFFGNHFFRANSTSDSVFTETPSVPL-----KNAIS 108
 870 PESGLQLRLNEKL--CTTAATELKKLD-----FKVSSSTNNLISTIPSDNLAAGDNTSS 922
 QY 109 LPTIGGSQALMLPLSPYTFNSKQES--FGPAKLRL--SCEPVMEKKAQKSSLLNGTV 165
 923 LG-----PPSMVHYDSQLDITTLFGKSSPLTESGGLSLSENNDSKLLSEGLM 972
 QY 166 HOGDTSWSSGSASOSS---QGRDSSHSSLSQYPPWPAEDMHPPTCELKGTSEE 222
 973 NSQESSWGKNVSTESGRLFKGKRAHGPALL-----TKDNALFKVISLSLTKNTKTSNN 1025
 QY 223 SLSD---LFGSLSLQLDLGSLDEVLN-----VMDKN 253
 1026 SATNRKTHIDGP--SLTIENSPSVWONILESDTEFKKVTPLIDRMMDKN 1074

RESULT 6
 RW1_HUMAN STANDARD; PRT; 1805 AA.
 ID Q92545; 2001 (Rel. 40, Created)
 AC 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RW1 protein (Fragment).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 RP TISSUE=Bone marrow;
 RC MEDLINE=9039502;
 RX Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawabayashi Y.,
 Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
 "Prediction of the coding sequences of unidentified human genes. VI.
 The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
 analysis of cDNA clones from cell line KG-1 and brain.";
 DNA Res. 3:321-329(1996).
 RT
 RL


```
D8 SOETPSTEE$RAONVFGQDNEDSDNLFGETESSVSNNANTPSIP-----TPVDN 138  
QY 128 FNSK-----QESEFPAKLPRISCEPFVMEKAOEKSLLENGTVHQDTSWGSSGSA 178  
    :| |      |||||       :| |:|   :| | :|  
Db 139 ENNKPAIKEDSTIQDSNGDVK---NMEDVKIOKEEFP----ENNTVIEG-----Y 181  
  
QY 179 SSSQGHDSSHSSLSEQYPDWPAEDMFDP--TPECLIGKTKSESLSLDTGLSLLOL 236  
    :| | :| ::| |     |  :| |:| :| | | |||         | :| :  
Db 182 KEESQ-PDEMTKEMDEVVEED---DEDDQPMSLPDNSIFGDTKSESKQLGNMFSVANTPS 237  
  
QY 237 DLGPSLDDEVLVNMDKNK 254  
    ::| | |:::||:  
Db 238 EIPDAHKAEOEDIIEKTE 255
```

```

RESULT 9
FRQ_CREP STANDARD; PRT; 1015 AA.
ID FRQ_CREP
AC 000586;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Frequency clock protein.
GN FRQ.
OS Creopus spinulosus (Chromocrea spinulosa);
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OC NCBI_TaxID=110619;
OX
RN
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4896;
RX MEDLINE=97188515; PubMed=9037100;
RA Lewis M.T., Morgan L.W., Feldman J.F.;
RT "Analysis of frequency (frq) clock gene homologs: evidence for a
RT helix-turn-helix transcription factor.";
RL Mol. Gen. Genet. 253:401-414(1997).
CC -1- FUNCTION: CIRCADIAN CLOCK COMPONENT INVOLVED IN THE GENERATION OF
CC BIOLOGICAL RHYTHMS, IN PARTICULAR IN RHYTHM STABILITY, PERIOD
CC LENGTH, AND TEMPERATURE COMPENSATION. BEHAVES AS A NEGATIVE
CC ELEMENT IN CIRCADIAN TRANSCRIPTIONAL LOOP (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FRQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U25850; AAA68072.1;
DR Biological rhythms; Transcription regulation; Nuclear protein.
KW 240 POLY-SER.
FT DOMAIN 356 POLY-GLN.
FT DOMAIN 443 POLY-PRO.
FT DOMAIN 443 POLY-PRO.
FT DOMAIN 584 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 892 ASF/GLU-RICH (ACIDIC).
FT DOMAIN 913 ASF/GLU-RICH (ACIDIC).
FT SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;
SQ
Query Match 7.28; Score 95.5; DB 1; Length 1015;
Best Local Similarity 19.68; Pred. No. 4.2; Indels 89; Gaps 10
Matches 55; Conservative 40; Mismatches 97;
OY 26 LSPDMISPLGDFRHTHIGKEGQ-----HDYFGDISFLQGYELLPGNOE 71
OY :||: :||:
DB 522 VTPDEVNAVVDLSAKFQLSPDGKIRWRGTDGTFKSSSSGDLQSRSPTDTEHTKG 581
DB :||: :||:
OY 72 KAHLGGPGHNEFFRANSTSDSVFTETPSPVLKNAISLPTIGSQALMLPLSPVTFNSK 131
OY | | | | | :||: |||
DB 582 DNHKRKTGH-----STGDSG-----SSGNLPKFGGPV-----SAS 614
DB :||: :||:
OY 132 QESFGPAKLPRLSCEPVMERKAQEKSSLENGTV-----HOGDTWSSGSSASQSSQ 183
OY ||| | :||: |||

```

```

RESULT 11
ZEP2_HUMAN
ID ZEP2_HUMAN STANDARD; PRT; 1833 AA.
AC P31629;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Human immunodeficiency virus type I enhancer-binding protein (HIV-Ep2).
GN HIVEP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91217105; PubMed=2022670;
RA Nomura N., Zhao M.-J., Nagase T., Maekawa T., Ishizaki R., Tabata S.,
RT Ishii S.;
RT "HIV-Ep2, a new member of the gene family encoding the human
RT immunodeficiency virus type I enhancer-binding protein. Comparison
RT with HIV-Ep1/PRD1-BF1/MBP-1.";
RL J. Biol. Chem. 266:8590-8594(1991).
RN [2]
RP SEQUENCE OF 1184-1323 FROM N.A.
RX MEDLINE=91062349; PubMed=2247438;
RA Rustgi A.K., Van'T Veer L.J., Bernards R.;
RT "Two genes encode factors with NF-kappa B- and H2TF1-like DNA-binding
RT properties.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8707-8710(1990).
CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
CC IN T-CELL ACTIVATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
CC -1- SIMILARITY: STRONG, TO HIVEP1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions.
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; M60119; AAB88218.1; -
DR EMBL; M61744; AAA86202.1; -
DR PIR; A39829; WMHUE2.
DR HSSP; P15832; 1BBO.
DR TRANSEAC; T00939; -
DR Genew; HGNC:4921; HIVEP2.
DR MIM; 143054; -
DR InterPro; IPR008822; Znf_C2H2.
DR Pfam; PF00096; Zf-C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 324 330
FT DOMAIN 337 369 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT FT 1286 1310 SER-RICH.
FT ZN_FING 1186 1208 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 1214 1238 C2H2-TYPE.
SQ SEQUENCE 1833 AA; 202128 MW; BAFDFA37BE075C23 CRC64;
Query Match
Best Local Similarity 7.1%; Score 93.5; DB 1; Length 1833;
Pred. No. 13;

```


DR MGD; MOD: 1007-1007

OS Schizosaccharomycetes pointue (*Schizosaccharomycetes*;
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

DR	EMBL: Z95396; CAB08763.1; -.
KM	Hypothetical protein; Transmembrane.
FT	TRANSMEM 209 229 POTENTIAL.
FT	TRANSMEM 241 261 POTENTIAL.
FT	TRANSMEM 267 287 POTENTIAL.
FT	TRANSMEM 328 348 POTENTIAL.
FT	TRANSMEM 361 381 POTENTIAL.
FT	TRANSMEM 387 407 POTENTIAL.
FT	TRANSMEM 917 937 POTENTIAL.
FT	TRANSMEM 975 995 POTENTIAL.
FT	TRANSMEM 997 1017 POTENTIAL.
FT	TRANSMEM 1021 1041 POTENTIAL.
FT	TRANSMEM 1066 1086 POTENTIAL.
FT	TRANSMEM 1275 1295 POTENTIAL.
SO	SEQUENCE 1337 AA; 150848 MW; A549BC8E0D08791 CRC64;

RESULT	14
ITF2_HUMAN	
ID	ITF2_HUMAN
AC	P15884; Q15439; Q15440;
DT	01-APR-1990 (Rel. 14, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Transcription factor 4 (Immunoglobulin transcription factor 2) (ITF-2)
DE	(SL3-3 enhancer factor 2) (SEF-2).
GN	TCF4 OR ITF2 OR SEF2.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxId-9606;

RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS SEF2-1A AND SEF2-1B).
 RC TISSUE=Thymocytes;
 RX MEDLINE=92015505; PubMed=1681116;
 RA Corneliussen B., Thornell A., Hallberg B., Grundstrom T.;
 RT "Helix-loop-helix transcriptional activators bind to a sequence in
 RL glucocorticoid response elements of retrovirus enhancers.";
 RN J. Virol. 65:6084-6093(1991).
 [2]
 RP SEQUENCE OF 46-667 FROM N.A. (ISOFORM SEF2-1B).
 RX MEDLINE=90175016; PubMed=2308860;
 RA Henthorn P., McCarrick-Walmsley R., Kadesch T.;
 RT "Sequence of the cDNA encoding ITF-2, a positive-acting transcription
 RL factor.";
 RN Nucleic Acids Res. 18:678-678(1990).
 [3]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=90140708; PubMed=2105528;
 RA Henthorn P., Kiledjian M., Kadesch T.;
 RT "Two distinct transcription factors that bind the immunoglobulin
 RL enhancer microE5/kappa 2 motif.";
 RN Science 247:467-470(1990).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IMMUNOGLOBULIN
 CC ENHANCER MU-E5/KES-MOTIF. BINDS TO THE E-BOX PRESENT IN THE
 CC SOMATOSTATIN RECEPTOR 2 INITIATOR ELEMENT (SSTR2-1NF) TO ACTIVATE
 CC TRANSCRIPTION (BY SIMILARITY). PREFERENTIALLY BINDS TO EITHER 5'-
 CC ACANNNGT-3', OR 5'-CCANNNG-3'.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. FORMS HOMO- OR HETERO-OLIGOMERS WITH MYOGENIN.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: SEF2-1A, SEF2-1B (SHOWN
 CC HERE) AND SEF2-1D; ARE PRODUCED BY ALTERNATIVE SPLICING. IN
 CC ADDITION, A MINI-EXON CODING FOR FOUR AMINO ACIDS MAY OR MAY NOT
 CC BE INCLUDED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT HEART, BRAIN, PLACENTA,
 CC SKELETAL MUSCLE AND TO A LESSER EXTENT IN THE LUNG. IN DEVELOPING
 CC EMBRYONIC TISSUES, EXPRESSION MOSTLY OCCURS IN THE BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; M74718; AAA60310.1; -;
 DR EMBL; M74719; AAA60311.1; -;
 DR EMBL; X52079; CAA36298.1; -;
 DR PIR; S08461; S08461.
 DR TRANSFAC; T00433; -;
 DR Genew; HGNC:11634; TCF4.
 DR MIM; 602272; -;
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS50888; HLH_2; 1.
 KW Transcription regulation; DNA-binding; Activator; Nuclear protein;
 KW Alternative splicing.
 FT DOMAIN 1 83
 FT
 FT DOMAIN 228 231
 FT DOMAIN 379 400
 FT DNA_BIND 564 576
 FT DOMAIN 577 618
 FT DOMAIN 619 642
 FT VARSPDIC 177 183
 FT VARSPDIC 545 545
 FT CONFLICT 46 49
 FT CONFLICT 334 334
 P -> S (IN REF. 2)
 ESSENTIAL FOR MYOD INHIBITION (BY
 SIMILARITY).
 POLY-SER.
 LEUCINE-ZIPPER (POTENTIAL).
 BASIC DOMAIN.
 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 CLASS A SPECIFIC DOMAIN.
 PGLPSS -> RDIYNGKA (IN ISOFORM SEF2-1A).
 T -> TRSRS (IN ISOFORM SEF2-1D).
 TGSN -> EFGG (IN REF. 2).

Wed Apr 23 13:51:15 2003

us-09-806-276a-2.rsp

SEQUENCE 667 AA; 71308 MW; 53459FC7989D9487 CRC64;
Query Match Similarity 22.7%; Score 92; DB 1; Length 667;
Best Local Similarity 22.7%; Pred. No. 4.6; Indels 54; Gaps 13;
Matches 52; Conservative 34; Mismatches 89;
23 RDLSPPDMISPLGDERHTIHIGEGQHDVEGDISLQNYE--LLPGNOEKALHGOE-- 78
83 RDIGSHDNLSPPFVNSRIQSKT-ERGSYSSYGRSNIQCHQOQSLGDMGMNPGTISP 129
79 --PGHNEFFRANSTSDVETETPSPVLKNAISLPTIGSQALMLPLSPVTEN 195
142 TRPG-SQYQYSSNPNRRRPLHSSAMEVQTKVKKVPPGLPS-----SVYAPSASTADYN 185
130 SKQESFGPAKLPLRSCPEVMEKAEKSSLLLENGTVHQGDTSWGSSGSASQSSQ---GR 185
196 RDSFGY-PSSKPASTFP-----SSEFMQDG--HSSDPWSSSGMNQPGYAGMLGN 244
186 DSH---SSLSSEQYPDWPAEDMDHPTPCELIKGKTSESLDITGSL 231
245 SSHLPQSSSYCSLHP---HERLSYPS-----HSSADINSSL 277

RESULT 15
YA8D-SCHPO STANDARD: PRT: 899 AA.

AC 009778; Q09778; (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last chromosome I.
DE Hypothetical protein C22F3.13 in chromosome I.
GN SPAC22F3.13.
OS Schizosaccharomyces pombe (Fission yeast).
OC Schizosaccharomycetes; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
OX NCBI_TaxID=4896;
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell J., Hodgson G.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Jagsels K.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Squares D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Smeaton M., Walsh S.V., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymponiez B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Pohl T.M.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Dreano S., Gloux S., Moore K., Hurst S.M.,
RA Coffeau A., Aves S.J., Xiang Z., Hunt C., Moore K., Garzon A., Rhode G.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Garzon A., Rhode G.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Benito J.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Potashkin J.,
RA Cerutti L., Lowe T., McCormick W.R., Paulsen I., Nurse P.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.,
RT The genome sequence of Schizosaccharomyces pombe.
RL Nature 415:871-880(2002).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; 254285; CA91078.1; POTENTIAL.
CC EMBL; 254285; CA91078.1; POTENTIAL.
DR Hypothetical protein; Transmembrane.
KW TRANSMEM 70 POTENTIAL.
FT TRANSMEM 338 358 03D7E9CC8866ECAB CRC64;
FT SEQUENCE 899 AA; 103435 MW; 03D7E9CC8866ECAB CRC64;
Query Match Similarity 24.7%; Score 90.5; DB 1; Length 899;
Best Local Similarity 24.7%; Pred. No. 9.2; Indels 33; Gaps 8;
Matches 48; Conservative 28; Mismatches 85;
62 NYELLPGNOEKALHGFPGHNEFFRANSTSDVETETPSPVLKNAISLPTIGSQ 116
377 NOELL-STKSDGLGRHLAHSNFKYTAETELTDKSRWTRLDIAVAALCNLSNAYGIAE 435
117 ALMLPL-ISPVTNSKQESFGPAKLPR---LSCEPVMEKAEKSSLLLENGTVHQGDT 170
436 SYMDPFGKLPPTYEETSSATGLAYPNESHDIASEPFSISWPNPSI---SGSVHATP 492
171 SWGSSGSASQSSQGRDSHSSLSSEQYPDWPAEDMDHPTPCELIKGKTSESLDITGSL 231
493 FDKAQLSTNEDSYDNISHTGTSYSEGV-----SSIHVKGEGSNNLELTSESL 540
225 SDLTGSLSLQIDL 238
541 SSTNDYIRRLQIDL 554

Search completed: April 23, 2003, 11:01:45
Job time: 19.0943 secs

Wed Apr 23 13:51:16 2003

us-09-806-276a-2.rspt

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: April 23, 2003, 10:57:05 ; Search time 28.1066 Seconds
(without alignments)
1862.055 Million cell updates/sec

Title: US-09-806-276a-2
Perfect score: 1326
Sequence: 1 MPARTPIYLKANNKKKKF.....QLDGLPSLDELVLNVDKKNK 254

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriophage:
17: sp_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1326	100.0	254	4	Q9UK12	Q9UK12 homo sapien
2	1320	99.5	254	4	Q95353	Q95353 homo sapien
3	1319	99.5	254	4	Q9U0J0	Q9U0J0 homo sapien
4	1230	92.8	254	11	Q9CQC5	Q9CQC5 mus musculu
5	407.5	30.7	210	4	Q14613	Q14613 homo sapien
6	406.5	30.7	210	4	Q9UNSO	Q9UNSO homo sapien
7	406.5	30.7	210	4	Q9BS13	Q9BS13 homo sapien
8	216	16.3	409	11	Q9D8M1	Q9D8M1 mus musculu
9	216	16.3	409	11	Q9B513	Q9B513 mus musculu
10	131	9.9	148	4	Q8MB51	Q8MB51 homo sapien
11	130.5	9.8	1373	5	Q960Z6	Q960Z6 drosophila
12	130.5	9.8	1373	5	Q917F7	Q917F7 drosophila
13	124	9.4	150	11	Q9Q2T9	Q9Q2T9 mus musculu
14	122.5	9.2	150	11	Q9Z0X0	Q9Z0X0 mus musculu
15	108	8.1	619	4	Q9ULD9	Q9ULD9 homo sapien
16	104	7.8	2283	2	Q8VO99	Q8VO99 staphylococ

17	99	7.5	1160	4	Q9E014	Q9E014 mus musculu
18	98	7.4	762	5	Q9C0H9	Q9C0H9 homo sapien
19	98	7.4	1357	5	Q9W4M4	Q9W4M4 drosophila
20	98	7.4	2237	5	Q9V122	Q9V122 drosophila
21	97.5	7.4	1020	5	Q9W313	Q9W313 drosophila
22	97	7.3	556	5	Q9S5Q3	Q9S5Q3 drosophila
23	97	7.3	988	11	Q9E0H5	Q9E0H5 rattus norv
24	97	7.3	1914	3	Q8X010	Q8X010 neurospora
25	97	7.3	2271	16	Q990Y4	Q990Y4 staphylococ
26	96.5	7.3	1195	5	Q8T433	Q8T433 drosophila
27	96.5	7.3	1447	10	Q64473	Q64473 arabidopsis
28	95.5	7.2	1955	3	Q9P8U6	Q9P8U6 neurospora
29	94.5	7.1	383	5	Q96ME2	Q96ME2 geodia cydo
30	94.5	7.1	912	4	Q96ME2	Q96ME2 homo sapien
31	94.5	7.1	1258	4	Q9P2A7	Q9P2A7 homo sapien
32	94	7.1	416	11	Q990T2	Q990T2 mus musculu
33	94	7.1	764	10	Q9SCY9	Q9SCY9 arabidopsis
34	94	7.1	3012	4	Q9Y3N6	Q9Y3N6 mus musculu
35	93.5	7.1	1019	11	Q9ESL6	Q9ESL6 mus musculu
36	93.5	7.1	1779	11	Q9VND9	Q9VND9 mus sapien
37	93.5	7.1	2446	4	Q9NS05	Q9NS05 homo sapien
38	93.5	7.1	2500	4	Q02646	Q02646 homo sapien
39	93	7.0	606	4	Q9NP00	Q9NP00 caenorhabdi
40	92.5	7.0	1144	5	P91389	P91389 xenopus lae
41	92	6.9	284	13	Q918H0	Q918H0 human immun
42	92	6.9	490	15	Q79356	Q79356 human immun
43	92	6.9	796	5	Q9TVE4	Q9TVE4 caenorhabdi
44	92	6.9	817	5	Q9XTE3	Q9XTE3 caenorhabdi
45	92	6.9	1042	16	Q8XAM9	Q8XAM9 escherichia

ALIGNMENTS

RESULT 1	ID	PRELIMINARY:	PRT:	254 AA.
Q9UK12	Q9UK12	Q9UK12		
AC	01-MAY-2000 (TREMREL. 13, Created)			
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)			
DT	01-MAR-2002 (TREMREL. 20, Last annotation update)			
DE	CRIB-containing BORG2 protein (Hypothetical 27.7 kDa protein).			
GN	BORG2 OR DKFZP434A0530.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=HEART;			
RC	MEDLINE=99421943; PubMed=10490598;			
RX	Joberty G., Perlungher R.R., Macara I.G.;			
RA	"The BORGs, a new family of Cdc42 and Tc10 GTPase-interacting			
RT	proteins."			
RT	Mol. Cell. Biol. 19:6585-6597(1999).			
RL				
RP	SEQUENCE FROM N.A.			
RP	TISSUE=TESTIS;			
RC	Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;			
RA	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RL	[3]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=LUNG;			
RC	Strausberg R.;			
RA	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AF164118; AAD48815.1;			
DR	EMBL; AL136842; CAB66776.1;			
DR	EMBL; BC019270; AAH19270.1;			
DR	InterPro; IPR000095; PAKbox/Rhodning.			
DR	SMART; SM00285; PBD; 1.			
KW	Hypothetical protein.			
SO	SEQUENCE 254 AA; 27678 MW; C31CDBE151C161ED CRC64;			

Query Match
Best Local Similarity 100.0%; Score 1326; DB 4; Length 254;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
QY 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
DB 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
QY 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
DB 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
QY 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
DB 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
QY 241 SLDEVLNVMDKNK 254
DB 241 SLDEVLNVMDKNK 254

RESULT 2

ID 095353 PRELIMINARY; PRT: 254 AA.
AC 095353;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE M5E5-related protein.
GN UBI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-98204843; PubMed-9535835;
RA Alberts A.S., Bouguin N., Johnston L.H., Treisman R.;
RT "Analysis of RhoA-binding proteins reveals an interaction domain
RT conserved in heterotrimeric G protein beta subunits and the yeast
RL J. Biol. Chem. 273:8616-8622(1998).
DR EMBL; AF094521; AAC71773.1;
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 254 AA; 27738 MW; C31CDBEI51DA605D CRC64;

Query Match
Best Local Similarity 99.5%; Score 1320; DB 4; Length 254;
Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
QY 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
DB 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
QY 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
DB 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
QY 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
DB 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
QY 241 SLDEVLNVMDKNK 254

DB 241 SLDEVLNVMDKNK 254

RESULT 3

ID 090QJ0 PRELIMINARY; PRT: 254 AA.
AC 090QJ0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cdc42 effector protein 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PREGNANT UTERUS;
RX MEDLINE-21125609; PubMed-11035016;
RA Hirsch D.S., Pirone D.M., Burdello P.D.;
RT "A New Family of Cdc42 Effector Proteins, CEPS, Function in Fibroblast
RT and Epithelial Cell Shape Changes."
RL J. Biol. Chem. 276:875-883(2001).
DR Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104857; AAD16888.1;
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 254 AA; 27679 MW; DBC2886CDAC7B822 CRC64;

Query Match
Best Local Similarity 99.5%; Score 1319; DB 4; Length 254;
Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 1 MPAKTPYLYKAANNKKGKKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
QY 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
DB 61 GNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFETETPSVLKNAISLPTIGSSQALML 120
QY 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
DB 121 PLSPVTENSKQESFPAKLPRLSCEPVEEKAQEKSSILENGTVHOGDTISWSSGSASQ 180
QY 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
DB 181 SSQGRDSSSSLSSEQYPDWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGP 240
QY 241 SLDEVLNVMDKNK 254
DB 241 SLDEVLNVMDKNK 254

RESULT 4

ID 09CQC5 PRELIMINARY; PRT: 254 AA.
AC 09CQC5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE 3200001F04Rik protein (RIKEN CDNA 3200001F04 gene).
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD, AND EMBRYO;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.,
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK019447; BAB31723.1; -
DR EMBL; AK012309; BAB28155.1; -
DR EMBL; AK014281; BAB29241.1; -
DR EMBL; BC021409; AAH21409.1; -
DR MGD; MGI:1915275; 3200001F04Rik.
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 254 AA; 27686 MW; 9EB9991CF1EB678F CRC64;

Query Match 92.8%; Score 1230; DB 11; Length 254;
Best Local Similarity 92.5%; Pred. No. 2.7e-106;
Matches 235; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

OY 1 MPAKTPYILKANNKKGKFKLRDLISPDMSPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 1 MPAKTPYILKANNKKGKFKLRDLISPDMSPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
OY 61 GNYELLPGNQEKAHLGQFPGHNEFFRANSTSDSVFTETPPVYLKNAISLPTIGSQALML 120
DB 61 GNYELLPGNQEKAHLGQFPGHNEFFRANSTSDSVFTETPPVYLKNAISLPTIGSQALML 120
OY 121 PLISPVTENSKQESFGPAKLRLSCPEVMEKEAQEKSSLENGTVHOGDTISWSSGSSAQ 180
DB 121 PLISPVTENSKQESFGPAKLRLSCPEVMEKEAQEKSSLENGTVHOGDTISWSSGSSAQ 180
OY 181 SSQGRDSSSSLSSEQYPPDPAEDMFDPHPTPCCELLIKGKTSEESLSLTLQIDLGP 240
DB 181 SSQGRDSSSSLSSEQYPPDPAEDMFDPHPTPCCELLIKGKTSEESLSLTLQIDLGP 240
OY 241 SLDELVLNMDKNK 254
DB 241 SLDELVLNMDKNK 254

RESULT 5
ID 014613 PRELIMINARY; PRT; 210 AA.
AC 014613;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 22.5 kda protein (Cdc42 effector protein 2).
GN BORG1.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97397562; Pubmed=9253601;
RA Guru S.C., Agarwal S.K., Manickam P., Olufemi S.-E., Crabtree J.S.,
RA Weisemann J.M., Kester M., Kim Y.S., Emmert-Buck M.R., Liotta L.A.,
RA Spiegel A.M., Boguski M., Roe B.A., Collins F.S., Burns A.L.,
RA Marx S.J., Chandrasekharappa S.C.,
RT "A transcript map for the 2.8-Mb region containing the multiple
RT endocrine neoplasia type 1 locus.";
RL Genome Res. 7:725-735(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=WHOLE BRAIN;
RX MEDLINE=99421943; Pubmed=10490598;
RA Joberty G., Perlungher R.R., Macara I.G.,
RT "The Borgs, a new family of Cdc42 and Tc10 GTPase-interacting
RT proteins.";
RL Mol. Cell. Biol. 19:6585-6597(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.,
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001436; AAB81206.1; -
DR EMBL; AF163840; AAD48784.1; -
DR EMBL; BC022337; AAH22337.1; -
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
KW Hypothetical protein.
SQ SEQUENCE 210 AA; 22483 MW; 7755100672FFB69C CRC64;

Query Match 30.7%; Score 407.5; DB 4; Length 210;
Best Local Similarity 41.2%; Pred. No. 6.6e-30;
Matches 107; Conservative 23; Mismatches 65; Indels 65; Gaps 8;

OY 1 MPAKTPYILKANNKKGKFKLRDLISPDMSPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 1 MPAKTPYILKANNKKGKFKLRDLISPDMSPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
OY 61 GNYELLPGNQEKAHLGQFPGHNEFFRANSTSDSVFTETPPVYLKNAISLPTIGSQAL 118
DB 61 GNYELLPGNQEKAHLGQFPGHNEFFRANSTSDSVFTETPPVYLKNAISLPTIGSQAL 118
OY 119 MLPLSPVTENSKQESFGPAKLRLSCPEVMEKEAQEKSSLENGTVHOGDTISWG 173
DB 119 MLPLSPVTENSKQESFGPAKLRLSCPEVMEKEAQEKSSLENGTVHOGDTISWG 173
OY 120 TLPTAQ-----APKPRRLHLE-TQPSDQEGSVDIWRIPETGSPNSGLT--- 164
DB 120 TLPTAQ-----APKPRRLHLE-TQPSDQEGSVDIWRIPETGSPNSGLT--- 164
OY 174 SSGSASOSSQGRDSSSSLSSEQYPPDPAEDMFDPHPTPCCELLIKGKTSEESLSLTLQ 233
DB 174 SSGSASOSSQGRDSSSSLSSEQYPPDPAEDMFDPHPTPCCELLIKGKTSEESLSLTLQ 233
OY 165 -----PESGAEEPF-----LSN-ASSLS 182
DB 165 -----PESGAEEPF-----LSN-ASSLS 182
OY 234 LQDLGPSLDELVLNMDKN 253
DB 234 LQDLGPSLDELVLNMDKN 253

RESULT 6
ID 09UNSO PRELIMINARY; PRT; 210 AA.
AC 09UNSO;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cdc42 effector protein 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
RX MEDLINE=99362714; PubMed=10430899;
RA Burbelo P.D., Snow D.M., Bahou W., Spiegel S.,
RT "MSE55, a Cdc42 effector protein, induces long cellular extensions in
RT fibroblasts.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9083-9088(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=21125609; PubMed=11035016;
RA Hirsch D.S., Pitone D.M., Burbelo P.D.,
RT "A New Family of Cdc42 Effector Proteins, CEPS, Function in Fibroblast
RT and Epithelial Cell Shape Changes.";
RL J. Biol. Chem. 276:875-883(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Thomas D.S., Burbelo P.D.,
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF098290; AAD16185.1;
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 210 AA; 22479 MW; 8255100667FFB936 CRC64;

Query Match 30.7%; Score 406.5; DB 4; Length 210;
Best Local Similarity 41.2%; Pred. No. 8.2e-30;
Matches 107; Conservative 23; Mismatches 65; Indels 65; Gaps 8;

QY 1 MPAKPIYLKANKKGRKFKLRDILSPDMISPLGDFRHTIHIGKQHDVFGDISFLQ 60
Db 1 MSTKYPIYLK-RGSRKGRKFKLRDLSSDMISPLGDFRHTIHIGSGGSDMFGDISFLQ 59
QY 61 GNYELLPGNOEKA--HIGQPGHNEFFRANSTSDSVETETPSVYLKNAISLPTIGSQAL 118
Db 60 GKFHLLPGTMEVEGEDTFLDFQFTRATVCGRELDPGSPPLKNAISLPIVIGGPQAL 119
QY 119 MLPLSPVTFNSKQESFGPAKLRLSCPEVMEKAQEKSS-----LLENGTVHQDTSWG 173
Db 120 TLPTAQ-----APPKPRHLLE-TPQSPQEGGSVDIWRIPETGSPNSGLT--- 164
QY 174 SSGSASQSSQGRDSSSSLSQYPPDPAEDMDHPTPCELIKGKTKSESLSDLTGSLLS 233
Db 165 -----PESGAEEPF-----LSN-ASSLLS 182
QY 234 LQDLGPSLLDEVLVNMDKN 253
Db 183 LHVDLGPSILDDVLIQIMDQD 202

RESULT 7
Q9BS13
ID Q9BS13; PRELIMINARY; PRT; 210 AA.
AC Q9BS13;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cdc42 effector protein 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVIX;
RA Strausberg R.,
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005406; AAH05406.1;
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 210 AA; 22456 MW; 775500E87C0FB69C CRC64;

Query Match 30.7%; Score 406.5; DB 4; Length 210;
Best Local Similarity 41.2%; Pred. No. 8.2e-30;

Matches 107; Conservative 22; Mismatches 66; Indels 65; Gaps 8;

QY 1 MPAKPIYLKANKKGRKFKLRDILSPDMISPLGDFRHTIHIGKQHDVFGDISFLQ 60
Db 1 MSTKYPIYLK-RGSRKGRKFKLRDLSSDMISPLGDFRHTIHIGSGGSDMFGDISFLQ 59
QY 61 GNYELLPGNOEKA--HIGQPGHNEFFRANSTSDSVETETPSVYLKNAISLPTIGSQAL 118
Db 60 GKFHLLPGTMEVEGEDTFLDFQFTRATVCGRELDPGSPPLKNAISLPIVIGGPQAL 119
QY 119 MLPLSPVTFNSKQESFGPAKLRLSCPEVMEKAQEKSS-----LLENGTVHQDTSWG 173
Db 120 TLPTAQ-----APPKPRHLLE-TPQSPQEGGSVDIWRIPETGSPNSGLT--- 164
QY 174 SSGSASQSSQGRDSSSSLSQYPPDPAEDMDHPTPCELIKGKTKSESLSDLTGSLLS 233
Db 165 -----PESGAEEPF-----LSN-ASSLLS 182
QY 234 LQDLGPSLLDEVLVNMDKN 253
Db 183 LHVDLGPSILDDVLIQIMDQD 202

RESULT 8
Q9D8M1
ID Q9D8M1; PRELIMINARY; PRT; 409 AA.
AC Q9D8M1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE 1810058K22R1K protein.
GN 1810058K22R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007896; BAB25335.1;
DR MGD; MGI:1917061; 1810058K22R1K.
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00285; PBD; 1.
SQ SEQUENCE 409 AA; 43064 MW; 7444D8005B1AAEB1 CRC64;

Query Match 16.3%; Score 216; DB 11; Length 409;
Best Local Similarity 32.5%; Pred. No. 1.1e-11;
Matches 82; Conservative 26; Mismatches 62; Indels 82; Gaps 14;

QY 26 LSPDMISPLGDFRHTIHIGKQHDVFGDISFL-----QGNVELLPGN-----QEKA 73

[illegible]

DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
KW ATP-binding; TRANSFERASE; 1.
SQ SEQUENCE 1337 AA; 147476 MW; 126509DF5A82724 CRC64;

Query Match
Best Local Similarity 24.3%; Score 130.5; DB 5; Length 1337;
Matches 63; Conservative 34; Mismatches 103; Indels 59; Gaps 13;
QY 1 MPAKPIYLAANKKGGKRRDLSPDMISPLGDFRHTIHKEGQHDVEGDISFL- 59
Db 462 LPSTRDSFSRSDHRSKRRLR-TEMISKPPQNDKFKHTGVGIDGA--TFGDIAFLG 515
QY 60 -QCNTELLPGNOEKALIGQFPGHNEFRANSTSDSVFTETP--SPVLKNAISLPTIGGS 115
Db 516 SSQNYNHVP-----KQIVTPYKPSEDI-EQTPLLPPTPTSPDSLQTAGS- 559
QY 116 QALMLPLSPVTENS--KQESFGPAKLPRLSCPEVMEKAKOEKSSLENGTV-----HQ 167
Db 560 -----YFPEGANGSGAMGTSMNPFIPSAEHTPKIATNGQSSDFEASGSTNPPFNR 612
QY 168 GDTSSWSSGASQSSQGRDSSSLSEQYPDWPAEDMFDHPTECELKGTKESESLSDL 227
Db 613 GDDDEL-EFGILNYGADGKSVHSET-----GM-----RPTSRSIVDDPEHYEISDD- 657
QY 228 TGSLSLQDLGSLDEY 246
Db 658 --ETADKLDGFSLLDEI 674

RESULT 12
Q917F7 PRELIMINARY; PRT; 1373 AA.
AC Q917F7; Q9V6K0; Q24316;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE Tyrosine-protein kinase PR2, last annotation update)
GN PR2 OR HD-11 OR CG3969.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Zhang Q., Chen L.X.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballew R.M., Basu A., Baxendale J., Andrews-Pfannkuch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhaktaroglu L., Beasley E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriter A., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Stimpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster";
RA Science 287:2185-2195(2000).
RP [2]
RC SEQUENCE OF 13-777 FROM N.A. (LONG ISOFORM).
RX STRAIN=OREGON-R; TISSUE=EMBRYO;
RA MEDLINE=94156202; PubMed=8112607;
RT Ito M., Matsui T., Taniguchi T., Chihara K.;
RT Drosophila melanogaster fibroblast growth factor receptor homolog.";
RN [3]
RP SEQUENCE OF 266-321 FROM N.A.
RX MEDLINE=98401146; PubMed=9731193;
RT Oates A.C., Wolberg P., Achen M.G., Wilks A.F.;
RT "Sampling the genomic pool of protein tyrosine kinase genes using the
RT polymerase chain reaction with genomic DNA.";
RT Biochem. Biophys. Res. Commun. 249:660-667(1998).
CC -1- TYROSINE PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC -1- ALTERNATIVE PRODUCTS.
CC -1- SIMILARITY: TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN.
CC EMBL: AE003819; AAF58423.2; ALT-SEQ.
CC EMBL: D17551; BAA04489.1; ALT-SEQ.
CC EMBL: AJ002909; CAA05744.1; -.
CC HSP; P13362; IEGK.
CC FLYBASE: FBgn0013955; PR2.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR000095; PAKDox/Rhodung.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00109; pkinase; 1.
CC PRODOM: PD000001; TYRKINASE.
CC SMART: SM00285; PBD; 1.
CC SMART: SM00220; S_TKC; 1.
CC SMART: SM00219; TYRK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS0011; PROTEIN_KINASE_TYR; 1.
CC TRANSFERASE; KINASE; Tyrosine-protein kinase; ATP-binding;
CC Alternative splicing.
CC NP_BIND 133 399
CC BINDING 139 147
CC ACT_SITE 164 164
CC DOMAIN 260 260
CC VARSPIC 505 539
CC CONFLICT 501 82
CC CONFLICT 501 501
CC CONFLICT 769 777
SQ SEQUENCE 1373 AA; 151489 MW; 4BC6597F764A8A12 CRC64;
Query Match
Best Local Similarity 9.8%; Score 130.5; DB 5; Length 1337;
Matches 63; Conservative 34; Mismatches 103; Indels 59; Gaps 13;
QY 1 MPAKPIYLAANKKGGKRRDLSPDMISPLGDFRHTIHKEGQHDVEGDISFL- 59
RA 1 MPAKPIYLAANKKGGKRRDLSPDMISPLGDFRHTIHKEGQHDVEGDISFL- 59

us-09-806-276a-2.rspt

wed Apr 23 13:51:16 2003

Db 479 LPSSTRDSFNRVSDHRSKRKLK---TEMISKPNDEKHTGHVIGDA--TFGDIAFLG 532
QY 60 -QGNVELLPNGNEKAHLGQFPNGNEFFRANSTSDSVFTEP---SPVLKNAISLPTIGS 115
Db 533 SSQNTNHP-----KQIVTPYKPSEDI-EQTPLLPPPTPTSPDLSQTASG- 576
QY 116 QALMLPLSPVTENS---KQESFGPAKLPLRSCPEYMEKEAKOESLLENGTV-----HQ 167
Db 577 -----YFPEGANGSGAMGTSMPPTFSAHTPKLATNGQSSDFASGSTNPFPPNR 629
QY 168 GDTSMGSSGSASQSGRDSHSSSLSEQYDPWPAEDMFDHPTPCELIKGKTSEESLSD 227
Db 630 GDEEL-EFGLHNGADGKSVHSET-----GW-----RPTRSIVDDPHYHEISDD- 674
QY 228 TGSLSLQDLGSLDEV 246
Db 675 --ETADKLDGSLDEI 691

RESULT 13
PRELIMINARY: PRT: 150 AA.

AC Q9QZT9: 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 20, Last annotation update)
DT 01-MAR-2002 (TREMblrel. 2010007002RIK protein) (Binder of RHO
DE CRIB-containing BORG3 protein (2010007002RIK protein))
DE GTPase 3).
GN BORG3 OR 2010007002RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=EMBRYO;
RC MEDLINE=99421943; PubMed=10490598;
RX MEDLINE=99421943; PubMed=10490598;
RA "The BORGs, a new family of Cdc42 and TCl0 GTPase-interacting
RT proteins." Mol. Cell. Biol. 19:6585-6597(1999).
RL [12]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
RC MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Boffelli D., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch R.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection." Nature 409:685-690(2001).
RL [3]
RN SEQUENCE FROM N.A.
RP Strausberg R.,
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164119; AAD48816.1;
DR EMBL; AK008154; BAB25498.1;
DR EMBL; BC006758; AA06758.1;
DR MGD; MGI:1929745; BORG3.

DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1. 15545 MW; 9AB9A49A9518B200 CRC64;
SQ SEQUENCE 150 AA; 15545 MW; 9AB9A49A9518B200 CRC64;

Query Match 9.4%; Score 124; DB 11; Length 150; Gaps 10;
Best Local Similarity 23.2%; Pred. No. 0.00096;
Matches 57; Conservative 14; Mismatches 39; Indels 136; Gaps 10;

QY 9 LKAANNKKGKFKLRDILSPDMISPLDGRHTIHIKREGQHDVFGDISFLQNGVELLP 49
Db 4 MKQLGPAQPKRRLDRLGALS--ISAPLDFRHTLHVGRG--DAFGDTSFL----- 128
QY 69 NOEKALGQFPNGNEFFRANSTSDSVFTEPSPVLKNAISLPTIGSQALMLPLSPVTF 68
Db 50 ---SRHGGGPP-----PEP-----GAP-----PVVAP----- 188
QY 129 NSKQESFGPAKLPLRSCPEYMEKEAKOESLLENGTVHQGDTSMGSSGSASQSGRDSH 83
Db 69 ---HSVAPPAPQ---PPVAV-----PSPAD-----PLSFHLDLGPMLDAVLG 107
QY 189 SSSLSQYDPWPAEDMFDHPTPCELIKGKTSEESLSDLTGSLSLQDLGSLDEVIN 248
Db 84 -----PSPAD-----PSPAD-----PSPAD-----PSPAD----- 107
QY 249 VMDKNK 254
Db 108 VMDAER 113

RESULT 14
PRELIMINARY: PRT: 150 AA.

AC Q9Z0X0: 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 20, Last annotation update)
DT 01-MAR-2002 (TREMblrel. 2010007002RIK protein) (Binder of RHO
DE CRIB-containing BORG3 protein (2010007002RIK protein))
DE GTPase 3).
GN BORG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21125609; PubMed=11035016;
RX Hirsch D.S., Pirone D.M., Burbelo P.D.,
RA "A New Family of Cdc42 Effector Proteins, CEPS, Function in Fibroblast
RT and Epithelial Cell Shape Changes." J. Biol. Chem. 276:875-883(2001).
RL EMBL; AF102773; AAD17906.1;
DR MGD; MGI:1929745; BORG3.
DR InterPro; IPR000095; PAKbox/Rhobndng.
DR SMART; SM00285; PBD; 1.
DR SEQUENCE 150 AA; 15630 MW; 8D26AAD58CC8B200 CRC64;

Query Match 9.2%; Score 122.5; DB 11; Length 150; Gaps 9;
Best Local Similarity 22.8%; Pred. No. 0.0013;
Matches 51; Conservative 12; Mismatches 28; Indels 133; Gaps 9;

QY 31 ISPLDGRHTIHIKREGQHDVFGDISFLQNGVELLPNGNEKAHLGQFPNGNEFFRANST 57
Db 23 ISAPLDFRHTLHVGRG--DAFGDTSFL-----SRHGGGPP----- 150
QY 91 SDSVFTEPSPVLKNAISLPTIGSQALMLPLSPVTFNSKQESFGPAKLPLRSCPEYME 83
Db 58 -----PEP-----GAP-----PVVAP-----HSVAPPAPQ---PPVAV 210
QY 151 EKAQKESLLENGTVHQGDTSMGSSGSASQSGRDSHSSLSLSEQYDPWPAEDMFDHPTP 86
Db 84 -----PSPAD-----PSPAD-----PSPAD-----PSPAD----- 254
QY 211 CELIKGKTSEESLSDLTGSLSLQDLGSLDEVINVMDKNK 254

Db 87 AD-----

PLSFHLDLGPMLDAVLGVMDAER 113

RESULT 15

Q9ULD9

ID Q9ULD9

AC Q9ULD9;

DT 01-MAY-2000

DT 01-MAY-2000

DT 01-MAY-2000

DE KIAA1281

GN KIAA1281

OS Homo sapiens

OC Eukaryota

OC Metazoa

OC Chordata

OC Vertebrata

OC Euteleostomi

OC Mammalia

OC Eutheria

OC Primates

OC Catarrhini

OC Homidae

OC Homo

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RX MEDLINE=20039619;

RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.,

RT "Prediction of the coding sequences of unidentified human genes. XV.

RT for large proteins in vitro."

RL DNA Res. 6:337-345(1999).

DR EMBL; AB033107; BAA86595.1; -.

FT NON_TER

FT SEQUENCE

619 AA; 68456 MW; 53C0243FD64AEAD0 CRC64;

Query Match

Best Local Similarity 8.1%; Score 108; DB 4; Length 619;

Matches 64; Conservative 34; Mismatches 113; Indels 56; Gaps 11;

QY 6 PIVKANNKRGKKKFLRDILSPDMISPLGDERHTIHIGKEGQHDVFGDISFLQNYEL 65

DB 42 PALVSLKDKKKKKKKRKLKDKGKGTGSPKM-DAK-----LGLLEDSKGASKD 87

QY 66 LPGNQEKALHGFPG-----HNEFFRANSTSDSVETET---PSPVLKNA--ISLPT 111

DB 88 LPHFLKDLNKNKNEGLANGLESQESRMASIKAEADKYTFEDNAPSPSISGASRLCST 147

QY 112 IGSSQALMLPLSPVTENSKQESFGPAKLRL-----SCEPYMEKAKQEKSS 158

DB 148 LVNGQAPMAP-LHVLTONGAESSAAKTSSPAYSDISDAADDGSDSRSEGMKRSKASSPSD 206

QY 159 LLENGTVHQDTSWSSGSSASQSSQGRDSSSSLSSEYPPDWPADDMFDHP---TPCELI 214

DB 207 IIS-----KDSVVKGHSSSTAQSSQKESHSPYHSHYDPIYSPSYM--HPGVGAPPAAGN 260

QY 215 KGKTK-----SEESLSDLTGSLSLQLD 237

DB 261 SGSTQGMKIKKKESEDAEKDKAEQLD 287

Search completed: April 23, 2003, 11:02:45
Job time : 31.1066 secs


```

OM protein - protein search, using OM
Run on:      April 23, 2003, 10:53:45 ; Search time 50.3484 Seconds
              (without alignments)
              619.298 Million cell updates/sec

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying criteria:

post-processing:	Minimum	Match	0%
	Maximum	Match	100%
		First 45 summaries	

```

A_Geneseq_101002:*
Database :
1:/SIDS2/gcgdata/geneseq/geneseqp-embl/A11980.DAT:*
1:/SIDS2/gcgdata/geneseq/geneseqp-embl/A11981.DAT:*
1:/SIDS2/gcgdata/geneseq/geneseqp-embl/A11982.DAT:*

```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1212	100.0	234	21	AAV92239	Human bone marrow-
2	1133.5	93.5	235	21	AAV93702	The kappa chain of
3	1133.5	93.5	235	21	AAV93729	The kappa chain of
4	1131	93.3	234	21	AAV93708	Light chain proteol
5	1127	93.0	234	21	AAV93733	The kappa chain of
6	1127	93.0	236	22	AAV93733	The kappa chain of
7	1116	92.1	236	23	ABG63490	Human gene 2-encod
8	1114.5	92.0	233	21	AAV93704	Human albumin fusi
9	1114.5	92.0	233	21	AAV93704	Human albumin fusi
10	1114.5	92.0	233	21	AAV93731	The kappa chain of

11	1108	91.4	236	23	AAU74299
12	1098	90.6	236	23	AAU74301
13	1084	89.4	234	15	AAR52951
14	1073	88.5	234	22	AAB36208
15	1055.5	87.1	349	12	AAR12128
16	1055.5	87.1	414	12	AAR13111
17	1055.5	87.1	414	12	AAR13018
18	1053	86.9	238	19	AAW83034
19	1053	86.9	238	21	AAAB14777
20	1053	86.9	238	21	AAW90927
21	1053	86.9	238	23	ABB74896
22	1053	86.9	238	23	ABB74942
23	1047	86.4	238	19	AAW83031
24	1047	86.4	238	21	AAAB14772
25	1047	86.4	238	21	AAW90922
26	1047	86.4	238	23	ABB74891
27	1047	86.4	238	23	ABB74937
28	1042.5	86.0	384	22	AAU14461
29	1042.5	86.0	384	22	AAU14462
30	1042.5	86.0	384	22	AAU14463
31	1042.5	86.0	384	22	AAU14464
32	1035	85.4	214	20	AAW08599
33	1033	85.2	224	22	AAW99395
34	1033	85.2	224	22	AAW99398
35	1033	85.2	224	22	AAW99399
36	1033	85.2	224	22	AAW75031
37	1033	85.2	224	22	AAW75032
38	1033	85.2	224	22	AAW75033
39	1033	85.2	224	22	AAW99372
40	1029	84.9	224	22	AAW75005
41	1029	84.9	238	19	AAW83032
42	1029	84.9	238	19	AAW83033
43	1029	84.9	238	21	AAAB14773
44	1029	84.9	238	21	AAAB14778
45	1029	84.9	238	21	AAAB14778

Anti-human AILIM m
Anti-human AILIM m
Human anti-IgE mAb
Human immune syste
1B1 IgG aberrant 1
1B1 IgG aberrant 1
1B1 IgG aberrant 1
Anti-Fas humanised
Humanised anti-Fas
Humanised HFE7A de
Humanised anti-Fas
Humanised anti-Fas
Anti-Fas humanised
Humanised anti-Fas
Humanised anti-Fas
Humanised anti-Fas
Humanised anti-Fas
Humanised anti-Fas
Humanised anti-Fas
Human novel protei
Human novel protei
Human novel protei
Human novel protei
Anti-human TNF- α p
Human interleukin
Human interleukin
Human interleukin
Human interleukin
Anti-IL8 monoclonal
Anti-IL8 monoclonal
Anti-IL8 monoclonal
gR0005 Humab kappa
Human Interleukin
Human IL8 monoclonal
Anti-Fas humanised
Anti-Fas humanised
Humanised anti-Fas
Humanised anti-Fas

RESULT 1
AAV92239 standard; Protein: 234 AA.
ID AAV92239
XX AC AAV92239;
DT XX 10-AUG-2000 (first entry)
DE XX Human bone marrow-derived serum protein 1.
XX Bone marrow-derived serum protein; immunoglobulin kappa light chain;
KW BMDSP-1; cytostatic; anti-HIV; antiasthmatic; antiatherosclerotic;
KW anti-inflammatory; antiarthritic; antisclerotic; hypotensive; antiviral;
KW anti-parasitic.
OS Homo sapiens.
XX Key location/Qualifiers
FH Modified-site 18 /note= "potential casein kinase II phosphorylation site"
FT Modified-site 34 /note= "potential casein kinase II phosphorylation site"
FT Domain 36..110 /label= immunoglobulin_domain
FT Modified-site 42 /note= "potential protein kinase C phosphorylation site"
FT Modified-site 72 /note= "potential protein kinase C phosphorylation site"
FT Modified-site 87 /note= "potential casein kinase II phosphorylation site"
FT Modified-site 96 /note= "potential casein kinase II phosphorylation site"
FT Modified-site 122 /note= "potential casein kinase II phosphorylation site"

0007-K-2000

99MO-ΠΣ23809
'6667-1.00-10

02-OCT-1998; 98US-016ECC

(INCY-) INCYTE PHARM INC

2 NC

...y MC, Guegler RJ, T.H. Davis

-PSDB; AAA09154

urified pol...

increased expression or activity of proteins preventing disorders

serum derived serum

English

serum proteins (BMDSP) 1

decreased expression of the *BMSP-1* gene in the presence of light chain.

are useful for treating associated with increased

Disorders include cancer of bone marrow.

... Crohn's disease ...

and vascular infections, parasitic diseases, and scleroderma.

ence 231 ,,

watch

Case	Similarity	Score	DB	Top1
234	100.08	1212	21	100.00

Mismatches: 0; Total:

SPATIAL SPACERATION

SPATIAL SPERATI C

00
00

-----G L P R F S G S G S G T D E T I T S B I M D N
-----S C Q 120

120

-----VAAPSVEIFPPSDEQIKSGTASWOT-----
-----ZGNSQ 180

180 QSGNQLDVALV

234

1

ID	Accession	Protein	Length	Score	DB	Length	Score	DB
181	ESVTEQDSKDSITYSLSTLTLSKADYEKKHYACEVTHOGLSSPVTKSFNREGC	234						
RESULT 2								
AA93702								
AA93702	standard; Protein; 235 AA.							
03-OCT-2000	(first entry)							
The kappa chain of immunoglobulin clone 4.1.1.								
Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system; hyperimmunity disorder; autoimmune disease; diabetes; graft rejection; proliferative disorder; cancer; immunodeficient disorder.								
Homo sapiens.								
WO200037504-A2.								
29-JUN-2000.								
23-DEC-1999;	99WO-US30895.							
23-DEC-1998;	98US-0113647.							
(PFIZ) PFIZER INC.								
(ABGE-) ABGENIX INC.								
Hanson DC, Neveu MJ, Mueller EE, Hanke JH, Gilman SC, Davis CG; Corvalan JR;								
WPI; 2000-442647/38.								
N-PSDB; AAA46865.								
Novel antibodies capable of binding cytotoxic T-lymphocyte antigen (CTLA)-4 containing specified heavy and light chain sequences, useful for treating, e.g. immune disorders								
Claim 3; Fig 1A; 157pp; English.								
The present sequence represents a kappa chain of an antibody of the invention. The antibody is directed cytotoxic T-lymphocyte antigen (CTLA)-4. Antibodies of the invention are composed of a heavy chain variable region, comprising a modified contiguous sequence from a FRL-FR3 sequence encoded by a human VH3-33 family gene. The modifications are contained in CDR1, CDR2 and/or framework regions. The antibodies may be used to inhibit CTLA-4 and down-regulate the immune system to treat hyperimmunity disorders (e.g. autoimmune disease, diabetes and graft rejection) and proliferative disorders (e.g. cancer). CTLA-4 stimulatory agents may be used to up-regulate immune system to up-regulate immunodeficient disorders.								
Sequence 235 AA;								
Query Match	93.5%;	Score 1133.5;	DB 21;	Length 235;				
Best Local Similarity	93.28;	Pred. No. 3.8e-62;						
Matches 219; Conservative	9;	Mismatches 6;	Indels 1;	Gaps 1;				
1 MEAPQQLFLLLMLPDTGTEIVLTQSPATLSISGERATISCRASQSV-SSYLAWYQOK 59								
1 MEAPQQLFLLLMLPDTGTEIVLTQSPATLSISGERATISCRASQSV-SSYLAWYQOK 59								
60 PGQAPRLILYDASNRATGIPRRFGSGSGTDFTLTISRLEPEDVALYYCOQYFTPTREG 119								
61 PGQAPRLILYDASNRATGIPRRFGSGSGTDFTLTISRLEPEDVALYYCOQYFTPTREG 119								
120 QGTRLEIKRTVAAPSVFIETPPSDEQLKSGTASVYVCLINNFYPRKRVQWKVDNALDLSGNS 179								
121 QGTRLEIKRTVAAPSVFIETPPSDEQLKSGTASVYVCLINNFYPRKRVQWKVDNALDLSGNS 179								

Db 181 QESVTEQDSKDYSLSTLTLTKADYEKKHVVACEVTHQGLSSPVTKSFNRGEC 235

RESULT 3
AAV93729 standard; Protein; 235 AA.

AAV93729;

03-OCT-2000 (first entry)

The kappa chain of immunoglobulin clone 4.1.1.

Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system; hyperimmunity disorder; autoimmune disease; diabetes; graft rejection; proliferative disorder; cancer; immunodeficient disorder.

Homo sapiens.

Key Location/Qualifiers
Peptide 1..20
/note= "signal peptide"

WO200037504-A2.

29-JUN-2000.

23-DEC-1999; 99WO-US30895.

23-DEC-1998; 98US-0113647.

(PEIZ) PEIZER INC.
(ABGE-) ABGENIX INC.

Hanson DC, Neveu MJ, Mueller EE, Hanke JH, Gilman SC, Davis CG; Corvalan JR;

WPI; 2000-442647/38.
N-PSDB; AAA46893.

Novel antibodies capable of binding cytotoxic T-lymphocyte antigen (CTLA)-4 containing specified heavy and light chain sequences, useful for treating, e.g. immune disorders -

Claim 3; Fig 22g; 157pp; English.

The present sequence represents a kappa chain of an antibody of the invention. The antibody is directed cytotoxic T-lymphocyte antigen (CTLA)-4. Antibodies of the invention are composed of a heavy chain variable region, comprising a modified contiguous sequence from a FRI-FR3 sequence encoded by a human VH3-33 family gene. The modifications are contained in CDR1, CDR2 and/or framework regions. The antibodies may be used to inhibit CTLA-4 and down-regulate the immune system to treat hyperimmunity disorders (e.g. autoimmune disease, diabetes and graft rejection) and proliferative disorders (e.g. cancer). CTLA-4 stimulatory agents may be used to up-regulate immune system to up-regulate immunodeficient disorders.

Sequence 235 AA;

Query Match 93.5%; Score 1133.5; DB 21; Length 235;
Best Local Similarity 93.2%; Pred. No. 3.8e-62;
Matches 219; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 MEAPQQLFLLLWLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 59
DB 1 MEAPQQLFLLLWLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 60
QY 60 GQAPRLLYDASNRAATGIPRFGSGSGTDFTLTISRLEPEDVALYCCOQYFTPTPTFG 119
DB 61 GQAPRLLYDASNRAATGIPRFGSGSGTDFTLTISRLEPEDFAVYCCOQYFTPTPTFG 120

QY 120 QGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLINFFPREAKVQWKNALQSGNS 179
DB 121 QGTKEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLINFFPREAKVQWKNALQSGNS 180

QY 180 QESVTEQDSKDYSLSTLTLTKADYEKKHVVACEVTHQGLSSPVTKSFNRGEC 234
DB 181 QESVTEQDSKDYSLSTLTLTKADYEKKHVVACEVTHQGLSSPVTKSFNRGEC 235

RESULT 4
AAO14066 standard; Protein; 234 AA.

AAO14066;

07-MAY-2002 (first entry)

light chain protein of the monoclonal antibody from clone JA.

HRIG; human rabies-immune globulin; monoclonal; virucide; heavy chain; human monoclonal rabies virus neutralising antibody; immunoglobulin; light chain; central nervous system; CNS; prophylactic therapy; clone JA.

Homo sapiens.

WO200188132-A2.

22-NOV-2001.

04-MAY-2001; 2001WO-US14468.

16-MAY-2000; 2000US-204518P.

(UYJE-) UNIV JEFFERSON THOMAS.

Hooper DC, Dietzschold B;

WPI; 2002-062381/08.
N-PSDB; AAK98702.

Novel isolated human monoclonal rabies virus neutralising antibody useful for treating individual exposed to rabies virus and for preventing spread of rabies virus to central nervous system -

Claim 4; Page 24-25; 25pp; English.

This sequence represents the light chain protein of the monoclonal antibody from clone JA. The invention relates to an isolated human monoclonal rabies virus neutralising antibody (virucide) derived from CDNA clones encoding the antibody heavy and light chains expressed in heterologous expression systems and purified away from deleterious contaminants. The invention provides a fused gene encoding a chimeric immunoglobulin light chain and a fused gene encoding a chimeric immunoglobulin heavy chain. The antibody of the invention is useful for treating an individual exposed to a rabies virus by administering to the individual a therapeutically effective amount of the antibody, and preventing a spread of the rabies virus to the central nervous system (CNS). The antibody of the invention provides a safe and efficacious post-exposure prophylactic therapy for individuals exposed to a rabies virus.

Sequence 234 AA;

Query Match 93.3%; Score 1131; DB 23; Length 234;
Best Local Similarity 92.7%; Pred. No. 5.4e-62;
Matches 217; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MEAPQQLFLLLWLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOK 60
DB 1 MEAPQQLFLLLWLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOK 60
QY 61 GQAPRLLYDASNRAATGIPRFGSGSGTDFTLTISRLEPEDVALYCCOQYFTPTPTFG 120
DB 61 GQAPRLLYDASNRAATGIPRFGSGSGTDFTLTISRLEPEDFAVYCCOQYFTPTPTFG 120

QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
||:::|||||
Db 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
QY 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
|||||
Db 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 5
AA93708
ID AAY93708 standard; Protein; 234 AA.
XX
AC AAY93708;
XX
DT 03-OCT-2000 (first entry)
XX
DE The kappa chain of immunoglobulin clone 6.1.1.
XX
KW Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system;
KW hyperimmunity disorder; autoimmune disease; diabetes; graft rejection;
KW proliferative disorder; cancer; immunodeficient disorder.
XX
OS Homo sapiens.
XX
PN WO200037504-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US30895.
XX
PR 23-DEC-1998; 98US-0113647.
XX
PA (PF12) PFIZER INC.
PA (ABGE-) ABGENIX INC.
XX
PI Hanson DC, Neveu MJ, Mueller EE, Hanke JH, Gilman SC, Davis CG;
PI Corvalan JR;
XX
DR WPI; 2000-442647/38.
DR N-PSDB; AAA46871.
XX
PT Novel antibodies capable of binding cytotoxic T-lymphocyte antigen
PT (CTLA)-4 containing specified heavy and light chain sequences, useful
PT for treating, e.g. immune disorders -
XX
PS Claim 3; Fig 1D; 157pp; English.
XX
CC The present sequence represents a kappa chain of an antibody of the
CC invention. The antibody is directed cytotoxic T-lymphocyte antigen
CC (CTLA)-4. Antibodies of the invention are composed of a heavy chain
CC variable region, comprising a modified contiguous sequence from a
CC FRI-FR3 sequence encoded by a human VH3-33 family gene. The
CC modifications are contained in CDR1, CDR2 and/or framework regions.
CC The antibodies may be used to inhibit CTLA-4 and down-regulate the
CC immune system to treat hyperimmunity disorders (e.g. autoimmune
CC disease, diabetes and graft rejection) and proliferative disorders
CC (e.g. cancer). CTLA-4 stimulatory agents may be used to up-regulate
CC immune system to up-regulate immunodeficient disorders.
XX
SQ Sequence 234 AA;

Query Match 93.0%; Score 1127; DB 21; Length 234;
Best Local Similarity 92.7%; Pred. No. 9.5e-62;
Matches 217; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MEAPAOQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKP 60
|||
Db 1 METPAQLFLLLMLPDTTGEIVLTQSPGTLSLSPGERATLSCRASQSVSYLAAYQOKP 60
QY 61 GQAPRLIYDASNRATGIPRPSGSGSGTDFLTISRLEPEDVALYYCQOYFTPTPTFGQ 120
|||||

Db 61 GQAPRLIYGVSRATGIPDRFSGSGGTDFTLTISRLEPEDRAVYYCQOYGISPTFGP 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
||:::|||||
Db 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
QY 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
|||||
Db 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 6
AA93733
ID AAY93733 standard; Protein; 234 AA.
XX
AC AAY93733;
XX
DT 03-OCT-2000 (first entry)
XX
DE The kappa chain of immunoglobulin clone 6.1.1.
XX
KW Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system;
KW hyperimmunity disorder; autoimmune disease; diabetes; graft rejection;
KW proliferative disorder; cancer; immunodeficient disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Peptide 1.20
FT /note= "signal peptide"
XX
PN WO200037504-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US30895.
XX
PR 23-DEC-1998; 98US-0113647.
XX
PA (PF12) PFIZER INC.
PA (ABGE-) ABGENIX INC.
XX
PI Hanson DC, Neveu MJ, Mueller EE, Hanke JH, Gilman SC, Davis CG;
PI Corvalan JR;
XX
DR WPI; 2000-442647/38.
DR N-PSDB; AAA46897.
XX
PT Novel antibodies capable of binding cytotoxic T-lymphocyte antigen
PT (CTLA)-4 containing specified heavy and light chain sequences, useful
PT for treating, e.g. immune disorders -
XX
PS Claim 3; Fig 22o; 157pp; English.
XX
CC The present sequence represents a kappa chain of an antibody of the
CC invention. The antibody is directed cytotoxic T-lymphocyte antigen
CC (CTLA)-4. Antibodies of the invention are composed of a heavy chain
CC variable region, comprising a modified contiguous sequence from a
CC FRI-FR3 sequence encoded by a human VH3-33 family gene. The
CC modifications are contained in CDR1, CDR2 and/or framework regions.
CC The antibodies may be used to inhibit CTLA-4 and down-regulate the
CC immune system to treat hyperimmunity disorders (e.g. autoimmune
CC disease, diabetes and graft rejection) and proliferative disorders
CC (e.g. cancer). CTLA-4 stimulatory agents may be used to up-regulate
CC immune system to up-regulate immunodeficient disorders.
XX
SQ Sequence 234 AA;

Query Match 93.0%; Score 1127; DB 21; Length 234;
Best Local Similarity 92.7%; Pred. No. 9.5e-62;
Matches 217; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MEAPAOQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKP 60

DB 1 MEXPQALLFLLLMLPDTGTEIVLTQSPGTLISLSPGERATLSCRASQSVSSYLAWYQOKP 60
OY 61 GOAPRLIYDASNRATGIPRPSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPEGQ 120
DB 61 GOAPRLIYGVSSSRATGIPRPSGSGSGTDFTLTISRLEPEDFAVYCCQYGISPTFGP 120
OY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
OY 181 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
RESULT 7
AAG71272 standard; Protein; 236 AA.
AAG71272;
AAG71272;
30-JUL-2001 (first entry)
Human gene 2-encoded secreted protein HTSER67, SEQ ID NO:121.
Human; secreted protein; proliferative disorder; cancer; chromosome 1;
foetal abnormality; developmental abnormality; haematopoietic disorder;
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
inflammation; allergy; neurological disorder; Alzheimer's disease;
Parkinson's disease; cognitive disorder; schizophrenia; asthma;
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
cardiovascular disorder; angiogenic disorder; kidney disorder;
gastrointestinal disorder; pregnancy-related disorder; tumour;
endocrine disorder; infection; wound healing; vulnery;
cell culture; chemotaxis; food additive;
binding partner identification.
Homo sapiens.
WO200132674-A1.
10-MAY-2001.
25-OCT-2000; 2000WO-US29360.
29-OCT-1999; 99US-0162211.
30-JUN-2000; 2000US-0215138.
(HUMA-) HUMAN GENOME SCI INC.
Ruben SM, Komatsoulis GA, Young PE, Moore PA;
WPI; 2001-291051/30.
N-PSDB; AAH31379.
New nucleic acid molecule encoding a human secreted protein, useful for
preventing, treating or ameliorating medical conditions such as
rheumatoid arthritis, Alzheimer's disease and microbial infections
Claim 11; Page 498-499; 581pp; English.
AAH31349-AAH31428 represent cDNAs corresponding to 26 human secreted
protein genes, and AAG71243-AAG71319 represent the proteins they encode.
AAG71320-AAG71403 represent human secreted protein fragments. The genes
and their corresponding secreted proteins are useful for preventing,
treating or ameliorating medical conditions, e.g., by protein or gene
therapy. Pathological conditions can be diagnosed by determining the
amount of the new protein in a sample or by determining the presence of
mutations in the new genes. Specific uses are described for each of the
52 genes, based on the tissues in which they are most highly expressed,
and include developing products for the diagnosis or treatment of
proliferative disorders, cancer, tumours, foetal and developmental

abnormalities, haematopoietic disorders, diseases of the immune system,
AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
allergies, neurological disorders (e.g., Alzheimer's disease,
Parkinson's disease), cognitive disorders, sepsis, diabetes, atherosclerosis,
skin disorders (e.g., psoriasis), angiogenic disorders, kidney disorders,
cardiovascular disorders, pregnancy-related disorders, endocrine
gastrointestinal disorders, The proteins can also be used to aid wound
healing and epithelial cell proliferation; to prevent skin aging due to
sunburn, to maintain organs before transplantation, for supporting cell
culture of primary tissues, to regenerate tissues, to identify their
culture of primary tissues, to regenerate tissues, and can be used
as a food additive or preservative to modify storage properties.
Antibodies specific for a protein of the invention can be used in
alleviating symptoms associated with the disorders mentioned above, and
in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
immunosorbent assay (ELISA). The present sequence represents a human
secreted protein of the invention.

Query Match 92.1%; Score 1116; DB 22; Length 236;
Best Local Similarity 92.8%; Pred. No. 4.5e-61;
Matches 219; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

OY 1 MEAPQALLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 60
DB 1 MEXPQALLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 60
OY 61 GOAPRLIYDASNRATGIPRPSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPEGQ 118
DB 61 GOAPRLIYXASXKRAATGIPRPSGSGSGTDFTLTISRLEPEDFAVYCCQYRXPXPTTF 120
OY 119 GOGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 178
DB 121 GOGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 180
OY 179 SOESTYEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 SOESTYEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 236

RESULT 8
ABG63490 standard; Protein; 236 AA.
ABG63490;
27-AUG-2002 (first entry)
Human albumin fusion protein #165.
Albumin fusion protein; therapeutic protein X; human albumin; HA;
human serum albumin; HSA; cancer; reproductive disorder;
digestive disorder; immune disorder; endocrine disorder;
haematopoietic disorder; neural disorder; connective disorder;
cytostatic; anti-infectivity; anti-inflammatory; anti-ulcer;
immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
neuroprotective; anti-Parkinsonian; antimicrobial; neuroleptic;
osteopathic; antiarthritic.
Homo sapiens.
Synthetic.
WO200177137-A1.
18-OCT-2001.
12-APR-2001; 2001WO-US11988.
12-APR-2000; 2000US-229358P.
25-APR-2000; 2000US-199384P.
21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2002-010886/01.

PT New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein
XX
PS Claim 1; Page 682; 2102pp; English.

CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, ulcerative colitis, digestive
CC disorders (e.g. Crohn's disease, reproductive disorders, immune disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.

CC
XX
SQ Sequence 236 AA;

Query Match 92.1%; Score 1116; DB 23; Length 236;
Best Local Similarity 92.8%; Pred. No. 4.5e-61;
Matches 219; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

QY 1 MEAPAQLLFLLLLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKP 60
DB 1 MEXPAQLFLLLLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKP 60
QY 61 GQAPRLIYDASNRAATGIPRFGSGSGSDFTLTISRLEPEDVALYYCQYFTTP--YTF 118
DB 61 GQAPRLIYDASNRAATGIPRFGSGSGSDFTLTISRLEPEDVALYYCQYFTTP--YTF 120
QY 119 GQTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 178
DB 121 GXGTKEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 180
QY 179 SQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 SQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 236

RESULT 9
AAV93704

ID AAV93704 standard; Protein; 233 AA.

AC AAV93704;

DT 03-OCT-2000 (first entry)

DE The kappa chain of immunoglobulin clone 4.8.1.

KW Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system;
KW hyperimmunity disorder; autoimmune disease; diabetes; graft rejection;
KW proliferative disorder; cancer; immunodeficient disorder.

OS Homo sapiens.

PN WO200037504-A2.

PD 29-JUN-2000.

PF 23-DEC-1999; 99WO-US30895.

PR 23-DEC-1998; 98US-0113647.

XX (PFIZ) PFIZER INC.
PA (ABGE-) ABGENIX INC.
XX
PI Hanson DC, Neveu MJ, Mueller EE, Hanke JH, Gilman SC, Davis CG;
PI Corvalan JR;
XX
DR WPI; 2000-442647/38.
DR N-PSDB; AAA46867.

PT Novel antibodies capable of binding cytotoxic T-lymphocyte antigen
XX (CTLA)-4 containing specified heavy and light chain sequences, useful
XX for treating, e.g. immune disorders
XX
PS Claim 3; Fig 1B; 157pp; English.

CC The present sequence represents a kappa chain of an antibody of the
CC invention. The antibody is directed cytotoxic T-lymphocyte antigen
CC (CTLA)-4. Antibodies of the invention are composed of a heavy chain
CC variable region, comprising a modified contiguous sequence from a
CC FRI-FR3 sequence encoded by a human VH3-33 family gene. The
CC modifications are contained in CDR1, CDR2 and/or framework regions.
CC The antibodies may be used to inhibit CTLA-4 and down-regulate the
CC immune system to treat hyperimmunity disorders (e.g. autoimmune
CC disease, diabetes and graft rejection) and proliferative disorders
CC (e.g. cancer). CTLA-4 stimulatory agents may be used to up-regulate
CC immune system to up-regulate immunodeficient disorders.

CC
XX
SQ Sequence 233 AA;

Query Match 92.0%; Score 1114.5; DB 21; Length 233;
Best Local Similarity 92.7%; Pred. No. 5.5e-61;
Matches 217; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 1 MEAPAQLLFLLLLPDPTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKP 60
DB 1 METPAQLFLLLLPDPTTGEIVLTQSPATLSLSPGERATLSCRISVS--SYLAWYQKP 59
QY 61 GQAPRLIYDASNRAATGIPRFGSGSGSDFTLTISRLEPEDVALYYCQYFTTPYTFGQ 120
DB 60 GQAPRLIYDASNRAATGIPRFGSGSGSDFTLTISRLEPEDVALYYCQYFTTPYTFGQ 119
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
DB 120 GTRVEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 179
QY 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 180 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 233

RESULT 10
AAV93731

ID AAV93731 standard; Protein; 233 AA.

AC AAV93731;

DT 03-OCT-2000 (first entry)

DE The kappa chain of immunoglobulin clone 4.8.1.

KW Cytotoxic T-lymphocyte antigen-4; CTLA-4; antibody; immune system;
KW hyperimmunity disorder; autoimmune disease; diabetes; graft rejection;
KW proliferative disorder; cancer; immunodeficient disorder.

OS Homo sapiens.

PN WO200037504-A2.

PF Key Location/Qualifiers

FT Peptide 1.20

/note= "signal peptide"

Db 181 SQESVTEQDSKSTYSLSSTLTLSKADYEKKHYACEVTHQGLSSPVTKEFNREGC 236

RESULT 12

AAU74301

ID AAU74301 standard; Protein; 236 AA.

AC AAU74301;

DE 12-MAR-2002 (first entry)

XX Anti-human ALLIM monoclonal antibody clone Jmab-139, light chain.

Human; antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
antiallergic; antiulcer; neuroprotective; antithyroid; vasotropic;
immunosuppressive; dermatological; antiinflammatory; hepatotropic;
activation inducible lymphocyte immunomodulatory molecule; ALLIM;
monoclonal antibody; allergy; rheumatoid arthritis; diabetes mellitus;
multiple sclerosis; autoimmune thyroiditis; psoriasis; hepatitis;
allergic contact-type dermatitis; chronic inflammatory dermatosis;
systemic lupus erythematosus; autoimmune disorder; inflammation;
graft versus host reaction; immune rejection; intestinal immunity;
ulcerative colitis; pneumonia; nephritis; vasculitis; pancreatitis.

OS Homo sapiens.

PN WO200187981-A2.

XX 22-NOV-2001.

PF 15-MAY-2001; 2001WO-JP04035.

PR 18-MAY-2000; 2000JP-0147116.

PR 30-MAR-2001; 2001JP-0099508.

PA (NISB) JAPAN TOBACCO INC.

PI Tsuji T, Tezuka K, Hori N;

DR WPI; 2002-075313/10.

DR N-PSDB; AAS99477.

XX New human monoclonal antibody that binds to activation inducible
lymphocyte immunomodulatory molecule, useful for treating rheumatoid
arthritis, multiple sclerosis and inflammation

PS Claim 30; Page 298-299; 300pp; English.

CC The invention relates to a novel human antibody (I), preferably a human
monoclonal antibody which binds to an activation inducible lymphocyte
immunomodulatory molecule (ALLIM). (I) is useful for modulating signal
transduction into a cell mediated by ALLIM, for modulating proliferation
of ALLIM-expressing cells, for inducing production of a cytokine from
ALLIM-expressing cells, and for inducing antibody-dependent cytotoxicity
against ALLIM-expressing cells and/or immune cytotoxicity or apoptosis of
prophylaxis of delayed type allergy. (I) is useful for preventing or
preventing various diseases associated with ALLIM-mediated
costimulatory transduction, and for inhibiting the onset and/or
advancement of the diseases. (I) is useful for suppression,
prevention and/or treatment of rheumatoid arthritis, multiple
sclerosis, autoimmune thyroiditis, allergic contact-type dermatitis,
chronic inflammatory dermatosis, systemic lupus erythematosus,
insulin-dependent diabetes mellitus, psoriasis, autoimmune or allergic
disorders, inflammation, graft versus host reaction, abnormal intestinal
immunity, immune rejection, disorders caused by abnormal intestinal
ulcerative colitis, pneumonia, hepatitis, nephritis, vasculitis, and
pancreatitis. (I) induces no serious immunorejection due to antigenicity
to human, i.e., human anti-mouse antigenicity (HAMA) in a host.
CC AAU74296-AAU74301 represent anti-human ALLIM monoclonal antibody amino
XX acid sequences of the invention.

SQ Sequence 236 AA:

Query Match

Best Local Similarity 90.6%; Score 1098; DB 23; Length 236;
Matches 216; Conservative 9; Mismatches 9; Indels 2; Gaps 2;

QY 1 MEAPAGLELLLMPLDPTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 59

Db 1 METPAQLLELLLMPLDPTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 60

QY 60 PGQAPRLIYDASNRATGIPPRFGSGSGTDEFTLTISRLPEEDVALYCCQYETTPY-TF 118

Db 61 PGQAPRLIFGASSRATGIPDRFGSGSGTDEFTLTISRLPEEDFAVYCCQYFGSSPMCSF 120

QY 119 GCGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 178

Db 121 GCGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 180

QY 179 SQESVTEQDSKSTYSLSSTLTLSKADYEKKHYACEVTHQGLSSPVTKEFNREGC 234

Db 181 SQESVTEQDSKSTYSLSSTLTLSKADYEKKHYACEVTHQGLSSPVTKEFNREGC 236

RESULT 13

AAR52951

ID AAR52951 standard; Protein; 234 AA.

AC AAR52951;

DE 27-OCT-1994 (first entry)

XX Human anti-IgE MAb light chain.

KW Human IgE; CH4 region; triggers mediator release;

KW Mast cells; Monoclonal antibody; allergy.

OS Homo sapiens.

FE Key

FT Region

XX Location/Qualifiers
21..128
/label= light chain variable region

PN EP592230-A.

PD 13-APR-1994.

PF 07-OCT-1993; 93EP-0308006.

PR 07-OCT-1992; 92JP-0293800.

PA (SNOW) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Kobayashi F, Mizuno A, Morinaga T, Washida N;

PI Yoshida T;

DR WPI; 1994-120330/15.

DR N-PSDB; AAQ71872.

XX Human monoclonal anti-IgE peptide antibody - inhibits histamine
release from mast cells by allergen stimulation, useful for
preventing allergies

PS Claim 3; Page 12; 21pp; English.

CC AAR52951 shows the light chain of a human type anti-IgE peptide
monoclonal antibody which inhibits the signal transmission for
the release of chemical mediator from mast cells and basophils
stimulated with allergen. The antibody can be used for the
prophylaxis and the therapy of allergy.

SQ Sequence 234 AA:

Query Match

89.4%; Score 1084; DB 15; Length 234;

Wed Apr 23 13:51:07 2003

us-09-806-276a-1.1rag

Best Local Similarity 89.3%; Pred. No. 4e-59; Indels 0; Gaps 0;
Matches 209; Conservative 11; Mismatches 14;

OY 1 MEAPQALLFLILLMLPDTGELVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQORP 60
DB 1 MEAPQALLFLILLMLPDTGELVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQORP 60
OY 61 GOAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
DB 61 GOAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
OY 61 AQAPRLIYGASTRATGIPARFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
DB 61 AQAPRLIYGASTRATGIPARFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
OY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
OY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
OY 181 ESVTEQDSKDSYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 181 ESVTEQDSKDSYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234

RESULT 14
AAB36208
ID AAB36208 standard; protein; 234 AA.
XX AAB36208;
AC AAB36208;
DE 15-FEB-2001 (first entry)
XX Human immune system associated protein HISAP-6.
DE Human immune system associated protein; HISAP-6; immune disorder;
XX Human; immune system associated disease; cancer.
KW infection; autoimmune disease; cancer.
XX Homo sapiens.
OS
XX US6135941-A.
PN 24-OCT-2000.
XX 27-MAR-1998; 98US-0049672.
XX 27-MAR-1998; 98US-0049672.
XX 27-MAR-1998; 98US-0049672.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Yue H, Lal P, Corley NC, Guegler KJ, Baughn MR;
PI Hillman JL, Au-Young J;
XX WPI: 2001-030926/04.
DR N-PSDB; AAC66524.
XX New human immune system associated proteins (HISAP) and polynucleotides
PT encoding the HISAP, useful for diagnosing, treating or preventing
PT immune or cell proliferative disorders or infections
XX Claim 1; Column 59-60; 54pp; English.
PS
XX The present invention provides the coding and protein sequences for a
CC number of human immune system associated proteins (HISAPs). These can be
CC used in the diagnosis and treatment of various autoimmune disorders,
CC infections and cell proliferation diseases. The diseases include AIDS,
CC adult respiratory distress syndrome, anaemia, asthma, atherosclerosis,
CC Crohn's disease, irritable bowel syndrome, multiple sclerosis, myasthenia
CC gravis, osteoarthritis, rheumatoid arthritis, scleroderma, systemic lupus
CC erythematosus, arteriosclerosis, cirrhosis and cancer.
XX
SQ Sequence 234 AA;
Query Match 88.5%; Score 1073; DB 22; Length 234;
Best Local Similarity 88.0%; Pred. No. 1.9e-58; Indels 0; Gaps 0;
Matches 206; Conservative 12; Mismatches 16;

DB 1 MEAPQALLFLILLMLPDTGELVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQORP 60
OY 61 GOAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
DB 61 GOAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
OY 61 GQSPRLIYGASTRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
DB 61 GQSPRLIYGASTRATGIPRFGSGSGTDFLTISRLEPEDVALYCCOQYFTPTPTFGQ 120
OY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
OY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAKVQWKVDNALQSGNSQ 180
OY 181 ESVTEQDSKDSYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 181 ESVTEQDSKDSYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234

RESULT 15
AAR12128
ID AAR12128 standard; protein; 349 AA.
XX AAR12128;
AC AAR12128;
DE 01-AUG-1991 (first entry)
XX 1B1 IgG aberrant light chain with duplicated variable region.
DE 1B1 IgG aberrant light chain; variable region; duplication;
XX passive immunity; group B streptococci.
KW
XX Homo sapiens.
OS
XX Location/Qualifiers
FH Key 1..17
FT /label= leader peptide
FT /label= 180
FT /label= variable region
FT /label= "L'V 1"
FT /label= variable region
FT /label= "L'V 2"
FT /label= 244..345
FT /label= constant region
XX Region
XX WO9106305-A.
XX 16-MAY-1991.
XX 06-NOV-1990; 90WO-US06426.
XX 07-NOV-1989; 89US-0432700.
XX (BRIM) BRISTOL-MYERS SQUIB.
XX Shuford WW, Harris LJ, Rafi HV;
XX WPI: 1991-163947/22.
XX N-PSDB; AAQ11878.
XX Oligomeric immunoglobulin(s) with high avidity for antigen(s) -
PT formed by duplicating esp. variable region of light chain of IgG
PT class
XX Example 5; Fig 16; 104pp; English.
PS
XX This sequence is deduced from the cDNA clone 4B9-VK15 and includes
CC the amino acid sequence beyond the first stop codon. The clone is
CC incomplete, starting from the G of the ATG initiator codon, but the
CC initial Met is given. Antibody molecules of the invention can
CC include one or two aberrant light chains containing a duplicated
CC variable region, to produce heavier antibodies. These heavier
CC antibodies were found to have higher avidity than antibodies used to
CC just a single copy of the L'V region. The antibodies can be used to
CC treat disease, e.g. infection by Streptococcus agalactiae. They are

CC able to pass across the placenta.
CC See also AAQ11879 and AAQ11880.
xx
SQ Sequence 349 AA;

Query Match 87.13% Score 1055.5; DB 12; Length 349;
Best Local Similarity 63.0% Pred. No. 3.2e-57;
Matches 220; Conservative 4; Mismatches 10; Indels 115; Gaps 2;

```
QY 1 MEAPAOQLLELLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
DB 1 MEAPAOQLLELLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
QY 61 GQAPRLIYDASNRATGIPPRFSGSGGTDFLTISRLPEDEVALYYCQ 109
DB 61 GQAPRLIYDASNRATGIPPRFSGSGGTDFLTISRLPEDEVALYYCQ 109
QY 110 ----- 120
DB 121 GGTKEVEIKHTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 109
QY 110 ----- 180
DB 181 IYDASNRATGIPPRFSGSGGTDFLTISRLPEDEVALYYCQHRDNMPGATFGGTYE 125
QY 126 IKRTVAAPSVFIFPPSDEQKSGTASVCLLNNFYPRFAKVQWKVDNALQSGNSQESYTE 240
DB 241 IKRTVAAPSVFIFPPSDEQKSGTASVCLLNNFYPRFAKVQWKVDNALQSGNSQESYTE 185
QY 186 QDSKDYSLSTLTLSKADYEKHKVYACEVTHQGLSSPYTKSFNRGEC 234
DB 301 QDSKDYSLSTLTLSKADYEKHKVYACEVTHQGLSSPYTKSFNRGEC 349
```

Search completed: April 23, 2003, 11:01:06
Job time : 52.3484 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:59:20 ; Search time 12.4672 Seconds
(without alignments)
552.246 Million cell updates/sec

Title: US-09-806-276A-1

Perfect score: 1212
Sequence: 1 MEAPAQLLFLLLMLPDTG.....EVTHQGLSSPVTKSFNRGEC 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1073	88.5	234	4	US-09-049-672A-6
2	1009.5	83.3	235	1	US-08-276-852-153
3	1009.5	83.3	235	1	US-08-899-575-153
4	1009.5	83.3	235	1	US-08-899-575-153
5	1009.5	83.3	235	5	PCT-US95-08743-153
6	1003	82.8	235	5	US-08-812-586-16
7	1001	82.6	240	4	US-09-301-593-36
8	994.5	82.1	241	2	US-07-916-098A-56
9	979.5	80.8	214	2	US-08-480-753-6
10	979.5	80.8	214	3	US-09-041-889-11
11	979.5	80.8	214	3	US-08-837-058-11
12	968.5	79.9	239	3	US-08-487-550-6
13	968	79.9	215	2	US-08-480-753-8
14	961	79.3	240	4	US-09-301-593-28
15	958	79.0	236	1	US-08-157-101A-5
16	955.5	78.8	233	4	US-09-485-737B-69
17	951.5	78.5	239	3	US-08-812-586-29
18	947.5	78.2	235	4	US-09-171-945-97
19	943	77.8	214	4	US-09-679-397-1
20	943	77.8	214	4	US-09-680-148-1
21	943	77.8	237	2	US-08-463-587A-25
22	943	77.8	237	2	US-08-463-667A-3
23	943	77.8	237	3	US-08-923-854-25
24	941	77.6	237	5	PCT-US91-09133-26
25	941	77.6	218	2	US-08-887-352B-13
26	941	77.6	218	3	US-08-466-151-9
27	941	77.6	218	4	US-09-109-207C-13

28	941	77.6	218	4	US-09-296-005-13	Sequence 13, Appl
29	941	77.6	218	4	US-08-466-163B-9	Sequence 9, Appl
30	939	77.5	232	1	US-08-704-744-80	Sequence 80, Appl
31	938	77.4	218	5	PCT-US96-13152-2	Sequence 2, Appl
32	936.5	77.3	233	2	US-07-934-373C-25	Sequence 25, Appl
33	936.5	77.3	233	3	US-08-437-642B-25	Sequence 25, Appl
34	936.5	77.3	233	3	US-08-146-206C-25	Sequence 25, Appl
35	936.5	77.3	233	5	PCT-US93-07832-25	Sequence 25, Appl
36	935	77.1	214	4	US-09-247-352-4	Sequence 4, Appl
37	935	77.1	214	4	US-09-466-635-4	Sequence 4, Appl
38	934	77.1	218	2	US-08-887-352B-15	Sequence 15, Appl
39	934	77.1	218	2	US-08-887-352B-17	Sequence 17, Appl
40	934	77.1	218	2	US-08-887-352B-19	Sequence 19, Appl
41	934	77.1	218	2	US-08-887-352B-24	Sequence 24, Appl
42	934	77.1	218	4	US-08-887-352B-15	Sequence 15, Appl
43	934	77.1	218	4	US-09-109-207C-15	Sequence 17, Appl
44	934	77.1	218	4	US-09-109-207C-17	Sequence 19, Appl
45	934	77.1	218	4	US-09-109-207C-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-049-672A-6
Sequence 6, Application US/09049672A
Patent No. 6135941
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Au-Young, Janice
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,672A
FILING DATE: HERewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ceirone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 234 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLSUCT01

CLONE: 2280869
US-09-049-672A-6

Query Match 88.5%; Score 1073; DB 4; Length 234;
Best Local Similarity 88.0%; Pred. No. 4e-80;
Matches 206; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 MEAPQLFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
DB 1 MEAPQLFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
QY 61 GOAPRLIYDASNRATGIPPRSGSGSGTFTLTISRLEPEDVALYCCQYFTPTFTG 120
DB 61 GOAPRLIYDASNRATGIPPRSGSGSGTFTLTISRLEPEDVALYCCQYFTPTFTG 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNS 180
DB 121 GTRVQIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNS 180
QY 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 ESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 2

US-08-276-852-153
Sequence 153, Application US/08276852
Patent No. 5652138

GENERAL INFORMATION:

APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 153:

SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-852-153

Query Match 83.3%; Score 1009.5; DB 1; Length 235;
Best Local Similarity 84.3%; Pred. No. 5.8e-75;
Matches 198; Conservative 11; Mismatches 25; Indels 1; Gaps 1;

QY 1 MEAPQLFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 59
DB 1 MGVTQVGLLLMLTDARCEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
QY 60 GOAPRLIYDASNRATGIPPRSGSGSGTFTLTISRLEPEDVALYCCQYFTPTFTG 119
DB 61 GOAPRLIYDASNRATGIPPRSGSGSGTFTLTISRLEPEDVALYCCQYFTPTFTG 119
QY 120 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNS 179
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNS 179
QY 180 QESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 QESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 3

US-08-899-575-153
Sequence 153, Application US/08899575
Patent No. 5770440

GENERAL INFORMATION:

APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 153:

SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-575-153

Query Match	83.3%	Score 1009.5	DB 1	Length 235
Best Local Similarity	84.3%	Pred. No. 5.8e-75		
Matches 198	Conservative 11	Mismatches 25	Indels 1	Gaps 1
QY	1	MEAPAOQLFLLLMLPDTTGEIVLTOSPATLSLSPGERATLSGRASQSV-SSYLAWYQOK	59	
DB	1	MGVPTQVYGLLLMLWLDARCEIVLTSPDGLTSLSPGERATFSCRSSHSIRSRVAMYQHK	60	
QY	60	PGOAPRLTIYDASNRAATGIPRFSGSGSGDFTLTISRLEEDVALYYCQYFTTPTFG	119	
DB	61	PGOAPRLVIHGVSNRASGISDRFSGSGGTDTLTITRVEPEDFALYYCQYVGA SSTFG	120	
QY	120	QGTRLLEIKRIVAAPSVFI FPPDEQLKSGTASVYCLLNFFYPREAKVQWKVDNALQSGNS	179	
DB	121	QGTKLERKRTVAPSVFI FPPDEQLKSGTASVYCLLNFFYPREAKVQWKVDNALQSGNS	180	
QY	180	QESVTEQDSKDSITYSLSTLTLSKADYEKKHVYACEVTHOGLSSPVTSENRGEC	234	
DB	181	QESVTEQDSKDSITYSLSTLTLSKADYEKKHVYACEVTHOGLSSPVTSENRGEC	235	

RESULT 4
US-08-899-575-153
; Sequence 153, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5804440lh Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-899-575-153

Query Match	83.3%	Score 1009.5	DB 1	Length 235
Best Local Similarity	84.3%	Pred. No. 5.8e-75		
Matches 198	Conservative 11	Mismatches 25	Indels 1	Gaps 1
QY	1	MEADPAQLFLLLMLPDTGTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOK	60	
Db	1	MGVPTQVGLLLMLMTDARCEIVLTQSPGTLSPGERATFSCRSSHRSIRSRVAMWYQHK	119	
QY	60	PGQAPRLIYDASNRATGIPRFESGSGSGDFTLTISRLEPEDVALYYCQQTPTPTFG	120	
Db	61	PGQAPRLYTHGVSNRASGISDRFSGSGSGTDFTLTITRVEPEDFALYYCQVYGASSYTFG	179	
QY	120	QGTRELEIKRTVAAPSVFIAPPSEDLKSGTASVYCLLNFFYPREAKVQMKVDNALQSGNS	180	
Db	121	QGTLELRKRTVAPASVFIAPPSEDLKSGTASVYCLLNFFYPREAKVQMKVDNALQSGNS	234	
QY	180	QESVTEQDSKDSYSLSTLTLSKADYEKKHYACAEVTHQGLSSPVTKSFNRGEC	235	
Db	181	QESVTEQDSKDSYSLSTLTLSKADYEKKHYACAEVTHQGLSSPVTKSFNRGEC		

```

RESULT 5
PCT-US95-08743-153
; Sequence 153, Application PC/TUS9508743
; GENERAL INFORMATION:
APPLICANT: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08743
FILING DATE: 11-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
INFORMATION FOR SEQ ID NO: 153:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-08743-153

Query Match      83.3%; Score 1009.5; DB 5; Length 235;
Best Local Similarity 84.3%; Pred. No. 5,8e-75;
Matches 198; Conservative 11; Mismatches 25; Indels 1; Gaps

114

```

Patent No. 6048704
GENERAL INFORMATION:
APPLICANT: Martin David Tilson
TITLE OF INVENTION: PURIFIED AND RECOMBINANT ANTIGENIC
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH ABDOMINAL AORTIC ANEURISM (AAA)
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,586
FILING DATE: 07-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53862-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-812-586-16

Query Match
Best Local Similarity 82.8%; Score 1003; DB 3; Length 235;
Matches 197; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 MEAPQLFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKP 60
DB 3 MRVPAQLGLLLMLPGARCAIRIAQSPSSLSASTGDRVTITCRASQGISNYLAWYQKP 62
QY 61 GQAPRLIYDASNRATGIPRFGSGSGTDTLTISRLEPEDVALYCCQYFTTPTFGQ 120
DB 63 GQAPRLIYDASNRATGIPRFGSGSGTDTLTISRLEPEDVALYCCQYFTTPTFGQ 122
QY 121 GTRLEIKRTVAAPSVFIPTPSDEQLKSGTASVCLLNNFYPRKAVQWKVDNALQSGNSQ 180
DB 123 GTRLEIKRTVAAPSVFIPTPSDEQLKSGTASVCLLNNFYPRKAVQWKVDNALQSGNSQ 180
QY 181 ESVTEDSKDSTYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 233
DB 183 ESVTEDSKDSTYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 235

RESULT 7
US-09-301-593-36
Sequence 36, Application US/09301593A
GENERAL INFORMATION:
APPLICANT: Park, John E.
APPLICANT: Garlin-Chesa, Pilar
APPLICANT: Bamberger, Uwe
APPLICANT: Leger, Olivier
APPLICANT: Saldanha, Jose W.
APPLICANT: Rettig, Wolfgang J.
TITLE OF INVENTION: FAP-specific Antibody with Improved Productibility
FILE REFERENCE: 0652.1890001

CURRENT APPLICATION NUMBER: US/09/301,593A
CURRENT FILING DATE: 1999-04-29
EARLIER APPLICATION NUMBER: EP 98107925.4
EARLIER FILING DATE: 1998-04-30
EARLIER APPLICATION NUMBER: US 60/086,049
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 36
LENGTH: 240
TYPE: PRT
ORGANISM: Homo sapiens
US-09-301-593-36

Query Match
Best Local Similarity 79.6%; Score 1001; DB 4; Length 240;
Matches 191; Conservative 25; Mismatches 18; Indels 6; Gaps 1;

QY 1 MEAPQLFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKP 60
DB 1 METDPLLLVLLWLPSSGDIWMTQSPDSLAVSLGERATINCKSSQSLYSRNQKNTLA 54
QY 55 WYQRPQAPRLIYDASNRATGIPRFGSGSGTDTLTISRLEPEDVALYCCQYFTT 114
DB 61 WYQRPQAPRLIYDASNRATGIPRFGSGSGTDTLTISRLEPEDVALYCCQYFTT 114
QY 115 PYTGGTRLEIKRTVAAPSVFIPTPSDEQLKSGTASVCLLNNFYPRKAVQWKVDNAL 174
DB 121 PLTEGGTRLEIKRTVAAPSVFIPTPSDEQLKSGTASVCLLNNFYPRKAVQWKVDNAL 174
QY 175 QSGNSQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 181 QSGNSQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 240

RESULT 8
US-07-916-098A-56
Sequence 56, Application US/07916098A
Patent No. 5871732
GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
APPLICANT: JOSEPH J.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08843
FILING DATE: No. 5871732ember 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732ember 27, 1990
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL

	MATCHES	190;	CONSERVATIVE	9;	MISMATCHES	12;	INDELS	1;	GAPS	1;
QY	24	LTQSPATLSLSPGERATLSCRASQSVS-SYLAWYQQKPGAPRLIYDASNRA	TGIPPRF	82						
DB	3	LTQSPGTLSLFPGERATLSCRASQRISTFLAWYQQKPGQSPRLIFDASTRA	PGLPDRF	62						
QY	83	SGSGSGETDTLTISRLEPEDVALYYCCOQYFTTPTFGGTRLEIKRTVAAPSVFI	FPPSD	142						
DB	63	SASWGSTDFTLTISRLEPEDFAVYYCQHVGSPMTFGGTKEIKRTVAAPSVFI	FPPSD	122						
QY	143	EQLKSQTASVCLLNFFYPREAKYQWKVDNALQSGNSQESVTEQDSKDSTYSLS	SLTTLS	202						
DB	123	EQLKSQTASVCLLNFFYPREAKYQWKVDNALQSGNSQESVTEQDSKDSTYSLS	SLTTLS	182						
QY	203	KADYEKKHYACEVTHQGSLSSPVTKSFNRGEC	234							
DB	183	KADYEKKHYACEVTHQGSLSSPVTKSFNRGEC	214							
RESULT 11										
US-08-837-058-11										
; Sequence 11, Application US/08837058										
; Patent No. 6074835										
; GENERAL INFORMATION:										
; APPLICANT: Braun, Jonathan										
; APPLICANT: Targan, Stephan R.										
; APPLICANT: Eggena, Mark										
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of										
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using										
; NUMBER OF SEQUENCES: 26										
; CORRESPONDENCE ADDRESS:										
; ADDRESSEE: Campbell & Flores LLP										
; STREET: 4370 La Jolla Village Drive, Suite 700										
; CITY: San Diego										
; STATE: California										
; COUNTRY: USA										
; ZIP: 92122										
; COMPUTER READABLE FORM:										
; MEDIUM TYPE: Floppy disk										
; COMPUTER: IBM PC compatible										
; OPERATING SYSTEM: PC-DOS/MS-DOS										
; SOFTWARE: Patentin Release #1.0, Version #1.25										
; CURRENT APPLICATION DATA:										
; APPLICATION NUMBER: US/08/837,058										
; FILING DATE:										
; CLASSIFICATION: 435										
; ATTORNEY/AGENT INFORMATION:										
; NAME: Campbell, Cathryn A.										
; REGISTRATION NUMBER: 31,815										
; REFERENCE/DOCKET NUMBER: P-PM 2438										
; TELECOMMUNICATION INFORMATION:										
; TELEPHONE: (619) 535-9001										
; TELEFAX: (619) 535-8949										
; INFORMATION FOR SEQ ID NO: 11:										
; SEQUENCE CHARACTERISTICS:										
; LENGTH: 214 amino acids										
; TYPE: amino acid										
; TOPOLOGY: linear										
; MOLECULE TYPE: protein										
US-08-837-058-11										
Query Match 80.8%; Score 979.5; DB 3; Length 214;										
Best Local Similarity 89.6%; Pred. No. 1.4e-72;										
Matches 190; Conservative 9; Mismatches 12; Indels 1; Gaps 1;										
QY	24	LTQSPATLSLSPGERATLSCRASQSVSYLAWYQQKPGAPRLIYDASNRA	TGIPPRF	82						
DB	3	LTQSPGTLSLFPGERATLSCRASQRISTFLAWYQQKPGQSPRLIFDASTRA	PGLPDRF	62						
QY	83	SGSGSGETDTLTISRLEPEDVALYYCCOQYFTTPTFGGTRLEIKRTVAAPSVFI	FPPSD	142						
DB	63	SASWGSTDFTLTISRLEPEDFAVYYCQHVGSPMTFGGTKEIKRTVAAPSVFI	FPPSD	122						
QY	143	EQLKSQTASVCLLNFFYPREAKYQWKVDNALQSGNSQESVTEQDSKDSTYSLS	SLTTLS	202						
DB	123	EQLKSQTASVCLLNFFYPREAKYQWKVDNALQSGNSQESVTEQDSKDSTYSLS	SLTTLS	182						
QY	203	KADYEKKHYACEVTHQGSLSSPVTKSFNRGEC	234							
DB	183	KADYEKKHYACEVTHQGSLSSPVTKSFNRGEC	214							
QY	63	SASWGSTDFTLTISRLEPEDFAVYYCQHVGSPMTFGGTKEIKRTVAAPSVFI	FPPSD	122						
DB	123	EQLKSQTASVCLLNFFYPREAKYQWKVDNALQSGNSQESVTEQDSKDSTYSLS	SLTTLS	182						
QY	83	SGSGSGETDTLTISRLEPEDVALYYCCOQYFTTPTFGGTRLEIKRTVAAPSVFI	FPPSD	142						
DB	63	SASWGSTDFTLTISRLEPEDFAVYYCQHVGSPMTFGGTKEIKRTVAAPSVFI	FPPSD	122						

QY	143	EQLKSGTASVCLLNFFYPREAKVQWKMVDNALQSGNSQESVTEQDSKDSTYSLSSTLLS	202
Db	123	EQLKSGTASVCLLNFFYPREAKVQWKMVDNALQSGNSQESVTEQDSKDSTYSLSSTLLS	182
QY	203	KADYEKKHVYACEVTHQGLSSPVTKSFNRGEC	234
Db	183	KADYEKKHVYACEVTHQGLSSPVTKSFNRGEC	214

RESULT 12
US-08-487-550-6
; Sequence 6, Application US/08487550
; Data set No. 612000

APPLICANT: Anderson, Darrell R.
 TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
 TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,
 TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
 TITLE OF INVENTION: IMMUNOSUPPRESSANTS".
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,550
FILING DATE: 07-JUN-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robln L.
REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 012712-131
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6666
TELEFAX: 703-836-3031

TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE: 1

SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-487-550-6

Query Match	79.9%;	Score 968.5;	DB 3;	Length 239;
Best Local Similarity	78.2%;	Pred. No. 1.3e-71;		
Matches 187; Conservative	23;	Mismatches 24;	Indels 5;	Cams 1

1 MEAPQLELLLLWLPDTGETVLTOSPATLSIPGERATLSCRASOV-----SSYIAN 55

Db

1 MSIPQQLGLLLCYGSSGEVMTOSPLSPITPGHPASTICRSSOSIKHSNCDTETSY 60

56 Y00KPG0APRLLYDASNRATGTIPBRESGSCSMBDEMT MIEDI ENDEUVAI VVZGCUWEMT 11

```

||||| | | | : | : ||||| : ||| : ||
61 YQOKDCDDPRT TVTGVNDCCNRNRCGCGCTTTT
db ----- -- -- -- -- -- -- -- -- --

```

01 LKXNTOYFRAHLLIIVNSNRUSV PDRISGSGAGIDFTLKRISAVEADVGVIYFCQGT RTP 12

11C *****

116 YTEGGGIRLEIKRTVAAPSVFIFFPPSDQLKSGTASVCLNNFFYPREAKYQWKYDALQ 17

D6 121 PTEGGGKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLNNFFYPREAKVQWKVDNALQ 18

176 SGNQSEVTEQDSKDYSLSTLTLSKADYEKKVYACEVTHQGLSSPYTKSFNRGEC 234

Db 181 SCNSQESTYEQDSKDYSLSTLTLSKADYEKKHVACEVTHQGLSSPYTKSINRGEC 239

```

; APPLICANT: Saldanha, Jose W.
; APPLICANT: Rettig, Wolfgang J.
; TITLE OF INVENTION: FAP-specific Antibody with Improved Productibility
; FILE REFERENCE: 0652.1890001
; CURRENT APPLICATION NUMBER: US/09/301,593A
; CURRENT FILING DATE: 1999-04-29
; EARLIER APPLICATION NUMBER: EP 98107925.4
; EARLIER FILING DATE: 1998-04-30
; EARLIER APPLICATION NUMBER: US 60/086,049
; EARLIER FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;
; LENGTH: 240
;
; TYPE: PR1
; ORGANISM: Homo sapiens
;
US-09-301-593-28

```

QY	1	MEAPAOLELLELLMLPDTIGELVIAQSFAISLSISFGEKATLSNNGEY	111
Db	1	MDSQAOVLMILLPLWVSGTCGDIWMSQSPSSLAVSVEKATMSCKSSQSLLYSRNQKNYLA	60
QY	55	WYQOKPGGAPRLLIYDASNRAITGIPRFSGSGSGTDFLTISRLEPEDVALYYCOOYFTT	114
Db	61	WFQOKPGQSPKLLIIFWASTRESGVPRDFTGSGFGTDFNLTISSVQAEDLAVYDCQOYFYSY	120
QY	115	PYTGGQTRLEIKRTVAAPSVFIPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL	174
Db	121	PLTFGAGTKLELKRITVAAPSVFIETPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL	180
QY	175	QSGNSQESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRRGEC	234
Db	181	QSGNSQESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRRGEC	240

RESULT 15
 US-08-157-101A-5
 ; Sequence 5, Application US/08157101A
 ; Patent No. 5808032
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: KURIHARA, TATSUYA
 ; APPLICANT: MATSUKURA, SHIGEKAZU
 ; APPLICANT: TSURUOKA, NOBUO
 ; APPLICANT: ARIMA, KENJI
 ; APPLICANT: NISHIHARA, TATSURO
 ; TITLE OF INVENTION: ANTI-HBS ANTIBODY GENES AND EXPRESSION
 ; TITLE OF INVENTION: PLASMIDS THEREFOR
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PILLSBURY, MADISON & SUTRO
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/157,101A
 ; FILING DATE: 05-APR-1994
 ; CLASSIFICATION: 530
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: TITUS, MARLANA K.
 ; REGISTRATION NUMBER: 35843
 ; REFERENCE/DOCKET NUMBER: 9437/204199
 ; TELECOMMUNICATION INFORMATION:
 ;

TELEPHONE: 202-861-3711
 TELEFAX: 202-822-0944
 TELEEX: 6714627 CUCH
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 236 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-157-101A-5

Query Match 79.0%; Score 958; DB 1; Length 236;
 Best Local Similarity 77.4%; Pred. No. 9e-71;
 Matches 181; Conservative 22; Mismatches 31; Indels 0; Gaps 0;

QY	1	MEAPQQLLELLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP	60
Db	3	MRVPAQLLGLLLWEPGARCIDIQTQSPSAMAASVGDRTITCRASQIGNYLVWFQOKP	62
QY	61	QAPRLIYDASNRTGIPPRFSGSGCTDFTLTISRLPEPDVALYCCQYFTTPYTFGQ	120
Db	63	GKVPKRLIYAASSLQSGVPSRFSGSGCTEFTLTISRLQPEDFATYYCLHHNNYPLSFGG	122
QY	121	GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLNNFYPREAKVQWKVDNALQSGNSQ	180
Db	123	GTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLNNFYPREAKVQWKVDNALQSGNSQ	182
QY	181	ESVTEQDSKDSSTYSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	234
Db	183	ESVTEQDSKDSSTYSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	236

Search completed: April 23, 2003, 11:03:52
 Job time : 14.4672 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:01:15 ; Search time 14.3852 Seconds
(without alignments)
1303.445 Million cell updates/sec

Title: US-09-806-276a-1
Perfect score: 1212
Sequence: 1 MEAPAQLLFLLLMLPDTG.....EVTGGLSSPYTKSFNRGEC 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1108	91.4	236	US-09-859-053-34	Sequence 34, Appl
2	1098	90.6	236	US-09-859-053-38	Sequence 38, Appl
3	999	82.4	240	US-09-799-514-8	Sequence 8, Appl
4	994	82.0	236	US-09-859-053-30	Sequence 30, Appl
5	991.5	81.8	239	US-09-249-011A-22	Sequence 22, Appl
6	991.5	81.8	239	US-09-924-340-8	Sequence 8, Appl
7	991.5	81.8	239	US-09-992-600A-8	Sequence 8, Appl
8	989	81.6	234	US-09-740-002-24	Sequence 24, Appl
9	986	81.4	212	US-10-006-593-118	Sequence 118, Appl
10	979	80.8	236	US-10-006-593-69	Sequence 69, Appl
11	968.5	79.9	239	US-10-124-905-6	Sequence 6, Appl
12	968.5	79.5	234	US-09-948-429B-6	Sequence 26, Appl
13	963.5	78.9	235	US-09-740-002-26	Sequence 150, Appl
14	956	78.8	234	US-09-800-729-150	Sequence 152, Appl
15	955.5	78.3	237	US-09-800-729-152	Sequence 10, Appl
16	949	78.3	237	US-10-020-786-10	Sequence 100, Appl
17	949	78.3	237	US-09-056-160B-100	Sequence 2, Appl
18	949	78.3	491	US-10-011-125-2	Sequence 2, Appl
19	948	78.2	218	US-09-925-179-67	Sequence 67, Appl

20	947.5	78.2	235	10	US-09-910-059-97	Sequence 97, Appl
21	945.5	78.0	239	10	US-09-825-012-9	Sequence 9, Appl
22	944.5	77.9	245	9	US-09-797-941A-6	Sequence 6, Appl
23	941	77.6	218	9	US-09-925-179-9	Sequence 9, Appl
24	941	77.6	218	10	US-09-802-077-9	Sequence 9, Appl
25	941	77.6	218	10	US-09-802-096-9	Sequence 13, Appl
26	941	77.6	218	10	US-09-920-171-13	Sequence 52, Appl
27	939.5	77.5	235	10	US-09-910-059-52	Sequence 2, Appl
28	938	77.4	218	10	US-09-917-410-2	Sequence 5, Appl
29	937	77.3	220	10	US-09-920-171-15	Sequence 15, Appl
30	934	77.1	218	10	US-09-920-171-17	Sequence 17, Appl
31	934	77.1	218	10	US-09-920-171-19	Sequence 19, Appl
32	934	77.1	218	10	US-09-920-171-19	Sequence 24, Appl
33	934	77.1	218	10	US-09-920-171-19	Sequence 24, Appl
34	933.5	77.0	235	10	US-09-910-059-99	Sequence 99, Appl
35	932	76.9	214	10	US-09-940-166A-2	Sequence 2, Appl
36	932	76.9	214	10	US-09-811-384-11	Sequence 11, Appl
37	932	76.9	237	9	US-10-020-786-8	Sequence 8, Appl
38	932	76.9	237	10	US-09-940-166A-6	Sequence 6, Appl
39	929.5	76.7	219	9	US-09-726-258-72	Sequence 72, Appl
40	929.5	76.7	242	9	US-09-726-258-51	Sequence 51, Appl
41	929.5	76.7	242	9	US-09-726-258-56	Sequence 56, Appl
42	929	76.7	214	9	US-09-875-221A-128	Sequence 128, Appl
43	929	76.7	214	10	US-09-949-559-128	Sequence 128, Appl
44	929	76.7	220	10	US-09-995-693-1	Sequence 1, Appl
45	928.5	76.6	242	9	US-09-726-258-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-859-053-34
Sequence 34, Application US/09859053
Patent No. US20020102658A1
GENERAL INFORMATION:
APPLICANT: Tezuka, Takashi
APPLICANT: Tezuka, Takashi
TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
TITLE OF INVENTION: PHARMACEUTICAL USE THEREOF
FILE REFERENCE: 06501-079001
CURRENT APPLICATION NUMBER: US/09/859, 053
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: JP 2001-99508
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: JP 2000-147116
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
US-09-859-053-34

QY	1	MEAPAQLLFLLLMLPDTG	91.4%;	Score 1108;	DB 10;	Length 236;
QY	1	MEAPAQLLFLLLMLPDTG	91.4%;	Pred. No. 4.6e-45;		
QY	1	MEAPAQLLFLLLMLPDTG	91.4%;	Best Local Similarity		
QY	1	MEAPAQLLFLLLMLPDTG	91.4%;	Matches 217;	Conservative	9; Mismatches 8; Indels 2; Gaps 2;
QY	60	PGQAPRLIYDASNRATGIPPRSSGSGTDTLTISRLPEDEVALYYCOQYFTTPY-TF				118
QY	61	PGQAPRLIYDASNRATGIPPRSSGSGTDTLTISRLPEDEVALYYCOQYFTTPY-TF				120
QY	119	GQGTLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLNNFYPREAKVQWKNALQSGN				178
QY	121	GQGTLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLNNFYPREAKVQWKNALQSGN				180

QY 179 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 181 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 236

RESULT 2
US-09-859-053-38

; Sequence 38, Application US/09859053
; Patent No. US20020102658A1
; GENERAL INFORMATION:
; APPLICANT: Tsuji, Takashi
; APPLICANT: Tezuka, Katsunari
; APPLICANT: Horii, No. US20020102658A1uaki
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
; FILE REFERENCE: 06501-079001
; CURRENT APPLICATION NUMBER: US/09/859, 053
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: JP 2001-99508
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: JP 2000-147116
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-859-053-38

Query Match
Best Local Similarity 90.6%; Score 1098; DB 10; Length 236;
Matches 216; Conservative 9; Mismatches 9; Indels 2; Gaps 2;

QY 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 59
DB 1 METPAQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 59
QY 60 PGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPY-TE 118
DB 61 PGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPY-TE 118
QY 119 GQGRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 178
DB 121 GQGRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGN 178
QY 179 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 181 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 236

RESULT 3

US-09-799-514-8
; Sequence 8, Application US/09799514
; Patent No. US20020065220A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: Immunoglobulin Superfamily Polynucleotides, Polypeptides, and An
; FILE REFERENCE: PT015PI
; CURRENT APPLICATION NUMBER: US/09/799, 514
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US00/23662
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/152,248
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-514-8

Query Match
Best Local Similarity 82.4%; Score 999; DB 10; Length 240;
Matches 191; Conservative 22; Mismatches 21; Indels 6; Gaps 1;

QY 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 54
DB 1 MVLQTVFISLLIMWISGAYGDIWTPQSPDSIAYSLGERATINCSSQTVLYSSDMKNYLA 60
QY 55 WYQKPGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPY 114
DB 61 WYQKPGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPY 114
QY 115 PYRFGQGRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNAL 174
DB 121 PYRFGQGRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNAL 174
QY 175 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 181 SQESVTEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 240

RESULT 4
US-09-859-053-30

; Sequence 30, Application US/09859053
; Patent No. US20020102658A1
; GENERAL INFORMATION:
; APPLICANT: Tsuji, Takashi
; APPLICANT: Tezuka, Katsunari
; APPLICANT: Horii, No. US20020102658A1uaki
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE ALLIM AND
; FILE REFERENCE: 06501-079001
; CURRENT APPLICATION NUMBER: US/09/859, 053
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: JP 2001-99508
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: JP 2000-147116
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-859-053-30

Query Match
Best Local Similarity 82.0%; Score 994; DB 10; Length 236;
Matches 190; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

QY 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 60
DB 3 MRVPAQLLGLLLWFPDSRCDIQMTQSPSSVSASVGDRTVITCRASQGISRLAWYQOK 62
QY 61 GQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPYFTGQ 120
DB 63 GQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPYFTGQ 122
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 180
DB 123 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 182
QY 181 ESVEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 234
DB 183 ESVEQDSKSTYSLSTLTLSKADYEKHKVYACEVTHOGLSSPVTKSFNRGEC 236

RESULT 5

US-09-249-011A-22
; Sequence 22, Application US/09249011A
; Patent No. US20020176855A1

GENERAL INFORMATION:
APPLICANT: CO, MAN SUNG
APPLICANT: VASQUEZ, MAXIMILIANO
APPLICANT: CARRENO, BEATRIZ
APPLICANT: CELMIKER, ABBIE CHERYL
APPLICANT: COLIINS, MARY
APPLICANT: GOLDMAN, SAMUEL
APPLICANT: GRAY, GARY S.
APPLICANT: KNIGHT, ANDREA
APPLICANT: O'HARA, DENISE
APPLICANT: RUP, BONITA
APPLICANT: VELDMAN, GEERTUIDA M.
TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
TITLE OF INVENTION: OF TREATMENT THEREWITH
FILE REFERENCE: 08702.0081-00000
CURRENT FILING DATE: 1999-02-12
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 22
LENGTH: 239
TYPE: PRT
ORGANISM: Mus sp.
US-09-249-011A-22

Query Match 81.8%; Score 991.5; DB 9; length 239;
Best Local Similarity 80.4%; Pred. No. 1.1e-39;
Matches 193; Conservative 21; Mismatches 19; Indels 7; Gaps 2;

QY 1 MEAPQQLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 54
1 MDSQAVLILLMLWSSGSDIVLTQSPDSLAVSLGERATISCSQSLLNSRTRENTLA 60
DB 55 WYQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 114
61 WYQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 120
QY 115 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNAL 174
121 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNAL 179
DB 175 QSGNSQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
180 QSGNSQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239

RESULT 6
US-09-924-340-8
Sequence 8, Application US/09924340
Publication No. US20030027248A1
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephanie
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91.US2.REG
CURRENT FILING DATE: 2001-08-06
CURRENT FILING DATE: 2001-08-06
PRIOR FILING DATE: 2001-07-13
PRIOR FILING DATE: 2001-07-13
PRIOR FILING DATE: 2001-06-29
PRIOR FILING DATE: 2001-06-29
PRIOR FILING DATE: 2001-06-15
PRIOR FILING DATE: 2001-06-15
PRIOR FILING DATE: 2001-05-25
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 8
LENGTH: 239
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL

LOCATION: 1..20
US-09-924-340-8

Query Match 81.8%; Score 991.5; DB 9; length 239;
Best Local Similarity 79.5%; Pred. No. 1.1e-39;
Matches 190; Conservative 22; Mismatches 22; Indels 5; Gaps 1;

QY 1 MEAPQQLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 55
1 MRLPAQLGLMLWSSGSDIVLTQSPDLFVTPGEPASISCRSSQSLHVGSNYLDW 60
DB 56 YQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115
61 YQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 120
QY 116 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQ 175
121 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQ 180
DB 176 SGNQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
181 SGNQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239

RESULT 7
US-09-992-600A-8
Sequence 8, Application US/0992600A
Publication No. US20030027161A1
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephanie
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91.US4.DIV
CURRENT FILING DATE: 2001-11-13
CURRENT FILING DATE: 2001-11-13
PRIOR FILING DATE: 2001-08-06
PRIOR FILING DATE: 2001-08-06
PRIOR FILING DATE: 2001-07-15
PRIOR FILING DATE: 2001-07-15
PRIOR FILING DATE: 2001-07-13
PRIOR FILING DATE: 2001-07-13
PRIOR FILING DATE: 2001-06-29
PRIOR FILING DATE: 2001-06-29
PRIOR FILING DATE: 2001-06-15
PRIOR FILING DATE: 2001-06-15
PRIOR FILING DATE: 2001-05-25
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 114
SOFTWARE: JPatent
SEQ ID NO 8
LENGTH: 239
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: 1..20
US-09-992-600A-8

Query Match 81.8%; Score 991.5; DB 9; length 239;
Best Local Similarity 79.5%; Pred. No. 1.1e-39;
Matches 190; Conservative 22; Mismatches 22; Indels 5; Gaps 1;

QY 1 MEAPQQLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 55
1 MRLPAQLGLMLWSSGSDIVLTQSPDLFVTPGEPASISCRSSQSLHVGSNYLDW 60
DB 56 YQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115
61 YQKPGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 120
QY 116 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQ 175
121 YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQ 180
DB 176 SGNQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
181 SGNQESVTEQDSKSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239

QY 176 SGNSESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 234
DB 181 SGNSESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 239

RESULT 8

US-09-740-002-24
; Sequence 24, Application US/09740002
; Patent No. US2002001798A1
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; APPLICANT: MORROW, PHILLIP
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
; FILE REFERENCE: 037003-0275759
; CURRENT APPLICATION NUMBER: US/09/740,002
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/335,697
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 08/488,376
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-002-24

Query Match
Best Local Similarity 81.6%; Score 989; DB 10; Length 234;
Matches 188; Conservative 19; Mismatches 27; Indels 0; Gaps 0;

QY 1 MEAPQQLFLLLWLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKP 60
DB 1 MEAPQQLGLLLWLGARCDIOMTQSPSSLSASVGDRTVITCRAGQRIASITLWYQHP 60
QY 61 GOAPRLIYDASNATGIPPRFSGSGGTDFLTISRLEPEDVALYCCQYFTTPTTPTFGQ 120
DB 61 GKAPKLLIYAGSNLHRGVPSRFSGSGGTDFLTITNSLQPEDFATYYCCQAYSTPTWTFGP 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQ 180
DB 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQ 180
QY 181 ESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 234
DB 181 ESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 234

RESULT 9

US-10-006-593-118
; Sequence 118, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 118

; LENGTH: 212
; TYPE: PRT
; ORGANISM: human
US-10-006-593-118

Query Match
Best Local Similarity 81.4%; Score 986; DB 9; Length 212;
Matches 193; Conservative 8; Mismatches 9; Indels 2; Gaps 2;

QY 24 LTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKPGQAPRLIYDASNATGIPPRF 82
DB 2 LTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKPGQAPRLIYDASNATGIPPRF 82
QY 83 SSGSGGTDFLTISRLEPEDVALYCCQYFTTPTTPTFGQTRLEIKRTVAAPSVFIFPPSD 142
DB 62 SSGSGGTDFLTISRLEPEDVALYCCQYFTTPTTPTFGQTRLEIKRTVAAPSVFIFPPSD 142
QY 143 EQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQSVTEODSKDSTYSLSTLTLS 202
DB 121 EQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQSVTEODSKDSTYSLSTLTLS 202
QY 203 KADYKHKYVACEVTHOGLSSPVTKSFNRGEC 234
DB 181 KADYKHKYVACEVTHOGLSSPVTKSFNRGEC 212

RESULT 10

US-10-006-593-69
; Sequence 69, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 69
; LENGTH: 236
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Humanized antibody light chain
US-10-006-593-69

Query Match
Best Local Similarity 80.8%; Score 979; DB 9; Length 236;
Matches 188; Conservative 17; Mismatches 29; Indels 0; Gaps 0;

QY 1 MEAPQQLFLLLWLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAAYQOKP 60
DB 3 MEAPQQLGLLLWLGARCDIOMTQSPSSLSASVGDRTVITCRAGQRIASITLWYQHP 60
QY 61 GOAPRLIYDASNATGIPPRFSGSGGTDFLTISRLEPEDVALYCCQYFTTPTTPTFGQ 120
DB 63 GKAPKLLIYAGSNLHRGVPSRFSGSGGTDFLTITNSLQPEDFATYYCCQAYSTPTWTFGP 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQ 182
DB 123 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKYQWKVDNALQSGNSQ 180
QY 181 ESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 234
DB 183 ESVTEODSKDSTYSLSTLTLSKADYKHKYVACEVTHOGLSSPVTKSFNRGEC 236

Wed Apr 23 13:51:08 2003

RESULT 11
US-10-124-905-6 Application US/10124905
; Sequence 6, Application US/10124905
; Patent No. US20020166136A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF, AS
; PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; IMMUNOSUPPRESSANTS"
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,905
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/383,916
; FILING DATE:
; APPLICATION NUMBER: US 08/487,550
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-10-124-905-6
; 79.9%; Score 968.5; DB 9; Length 239;
; Best Local Similarity 78.2%; Pred. No. 1.3e-38; Indels 5; Gaps 1;
; Matches 187; Conservative 23; Mismatches 24;
; QY 1 MEAPQQLLELLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 55
; 1 MSIPAOGLGLLLCVPGSSGEVMTQSPSLPTTPGEPASISCRSSQSLKHSNGDTFLSW 60
; DB 1 MSIPAOGLGLLLCVPGSSGEVMTQSPSLPTTPGEPASISCRSSQSLKHSNGDTFLSW 60
; QY 56 YQKPGQAPRLIYDASNRAIGIPPRFGSGSGTDFLTISRLEPEDVALYYCOQYFTTP 115
; 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 175
; 61 YQKPGQAPRLIYKVNRSRDSGVPDRFSGSGAGTDFTLKISAVEAEADVGVYFCGQGTPTP 120
; DB 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 180
; QY 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 180
; DB 121 PTFGGGTKEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 234
; QY 176 SGNSQSVTEODSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239
; DB 181 SGNSQSVTEODSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239

RESULT 12
US-09-948-429B-6 Application US/09948429B
; Sequence 6, Application US/09948429B

Patent No. US20020177689A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF, AS
; PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; IMMUNOSUPPRESSANTS"
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/948,429B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/383,916
; FILING DATE:
; APPLICATION NUMBER: US 08/487,550
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-948-429B-6
; 79.9%; Score 968.5; DB 9; Length 239;
; Best Local Similarity 78.2%; Pred. No. 1.3e-38; Indels 5; Gaps 1;
; Matches 187; Conservative 23; Mismatches 24;
; QY 1 MEAPQQLLELLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 55
; 1 MSIPAOGLGLLLCVPGSSGEVMTQSPSLPTTPGEPASISCRSSQSLKHSNGDTFLSW 60
; DB 1 MSIPAOGLGLLLCVPGSSGEVMTQSPSLPTTPGEPASISCRSSQSLKHSNGDTFLSW 60
; QY 56 YQKPGQAPRLIYDASNRAIGIPPRFGSGSGTDFLTISRLEPEDVALYYCOQYFTTP 115
; 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 175
; 61 YQKPGQAPRLIYKVNRSRDSGVPDRFSGSGAGTDFTLKISAVEAEADVGVYFCGQGTPTP 120
; DB 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 180
; QY 116 YTFGGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 180
; DB 121 PTFGGGTKEIKRTVAAPSVFIFPPSDEQLKSGTASVYCLINNEFYPREAKVQWKVDNALQ 234
; QY 176 SGNSQSVTEODSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239
; DB 181 SGNSQSVTEODSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 239

RESULT 13
US-09-740-002-26 Application US/09740002
; Sequence 26, Application US/09740002
; Patent No. US20020001798A1
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES

TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
FILE REFERENCE: MANUFACTURE AND THERAPEUTIC USE THEREOF
CURRENT APPLICATION NUMBER: 037003-0275759
CURRENT FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: US/09/740,002
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/335,697
PRIOR FILING DATE: 1999-06-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 26
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
US-09-740-002-26

Query Match
Best Local Similarity 79.5%; Score 963.5; DB 10; Length 234;
Matches 185; Conservative 24; Mismatches 23; Indels 1; Gaps 1;

QY 1 MEAPQLLELLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
DB 3 MRVPAQLLGLLMLRGARCDIQMTQSPSSLSASVGDRTVITCRASQSIASVYVYQOKP 60
QY 61 GQAPRLITYDASNRATGIPPRSGSGSGTDTLTISRLEPEDVALYYCOQYFTPTFGQ 120
DB 63 GKAPKVLIFASANLVSGVPSRSGSGTFTLTISNLPEDPAFTFCQDGTN-FSFGQ 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
DB 122 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
QY 181 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 182 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 14

US-09-800-729-150
Sequence 150, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
PRIOR APPLICATION NUMBER: 2001-03-08
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 150
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (120)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-800-729-150

Query Match
Best Local Similarity 78.9%; Score 956; DB 10; Length 234;
Matches 183; Conservative 20; Mismatches 31; Indels 0; Gaps 0;

QY 1 MEAPQLLELLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
DB 1 MRVPAQLLGLLMLRGARCDIQMTQSPSSLSASVGDRTVITCRASQSIASVYVYQOKP 60
QY 61 GQAPRLITYDASNRATGIPPRSGSGSGTDTLTISRLEPEDVALYYCOQYFTPTFGQ 120

DB 61 GKPKVIVFDGSIHTGVPSRSGSGTFTLTINNLPDDVATYSCQYNTFTPTFGX 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
DB 122 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
QY 181 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 182 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 15

US-09-800-729-152
Sequence 152, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
PRIOR APPLICATION NUMBER: 2001-03-08
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 152
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-152

Query Match

Best Local Similarity 78.8%; Score 955.5; DB 10; Length 235;
Matches 187; Conservative 16; Mismatches 30; Indels 1; Gaps 1;

QY 1 MEAPQLLELLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
DB 3 MRVPAQLLGLLMLRGARCDIQMTQSPSSLSASVGDRTVITCRASQSIASVYVYQOKP 60
QY 61 GQAPRLITYDASNRATGIPPRSGSGSGTDTLTISRLEPEDVALYYCOQYFTPTFGQ 120
DB 63 GKAPKVLIFASANLVSGVPSRSGSGTFTLTISNLPEDPAFTFCQDGTN-FSFGQ 120
QY 121 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
DB 122 GTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQ 120
QY 181 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 182 ESVTEQDSKDSSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

Search completed: April 23, 2003, 11:04:28
Job time : 15.3852 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:58:30 ; Search time 13.9057 Seconds
(without alignments)
1617.710 Million cell updates/sec

Title: US-09-806-276a-1

Perfect score: 1212

Sequence: 1 MEAPAQQLFLLLMLPDTG.....EVT HQGLSSPVTKSFNRGEC 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1018.5	84.0	215	2 JE0242	Ig kappa chain NIG
2	993.5	82.0	215	2 JE0244	Ig kappa chain NIG
3	974.5	80.4	215	2 A23746	Ig kappa chain V-I
4	956.5	78.9	215	2 JE0243	Ig kappa chain NIG
5	899	74.2	216	2 JE0241	Ig kappa chain Am3
6	806	66.5	240	2 S06084	Ig kappa chain pre
7	741	61.1	234	2 S01320	Ig kappa chain pre
8	740	61.1	234	2 S14237	Ig kappa chain pre
9	736.5	60.8	230	2 S33161	Ig kappa chain - s
10	726.5	59.9	225	2 S37484	Ig kappa chain - m
11	717	59.2	218	2 S68241	Ig kappa chain V r
12	712.5	58.8	219	2 S38865	Ig kappa chain - m
13	712	58.7	220	2 A31790	Ig kappa chain V r
14	711	58.7	218	2 JC5810	monoclonal antibod
15	706	58.3	214	2 S68212	Ig kappa chain (Ma
16	703.5	58.0	219	2 S52028	Ig kappa chain - m
17	703.5	58.0	235	2 S25058	Ig kappa chain - m
18	700.5	57.8	219	2 PC4203	Ig kappa chain (mo
19	700.5	57.8	219	2 S16112	Ig kappa chain V r
20	698.5	57.6	217	2 S42772	Ig kappa chain - m
21	693	57.2	210	2 A56169	Ig kappa chain V r
22	680.5	56.1	225	2 JI0029	Ig kappa chain pre
23	671	55.4	144	2 PL0106	Ig kappa chain pre
24	613.5	50.6	145	2 S20631	Ig kappa chain - h
25	613	50.6	178	2 PT0219	Ig kappa chain V-C
26	603.5	49.8	197	2 S29593	Ig kappa chain (WM
27	602	49.7	135	2 S52059	UC-kappa protein -
28	589.5	48.6	129	1 K3H0A	Ig kappa chain pre
29	589.5	48.6	229	2 A20969	Ig kappa chain pre

30	584.5	48.2	129	2 S49532	anti-Sm antibody V
31	583.5	48.1	129	1 K3H0H	Ig kappa chain pre
32	581.5	48.0	129	1 S46369	Ig light chain var
33	581.5	48.0	134	2 S38643	Ig kappa chain V r
34	580	47.9	128	2 S40379	Ig kappa chain V-J
35	577.5	47.6	128	2 S20636	Ig kappa chain V r
36	576	47.5	128	2 A56701	Ig kappa chain V r
37	573	47.3	128	1 K3H041	Ig kappa chain pre
38	572.5	47.2	129	2 S40363	Ig kappa chain - h
39	572.5	47.2	238	2 A49633	Ig kappa chain pre
40	571.5	47.2	129	2 A32274	Ig kappa chain pre
41	567	46.8	128	2 S40343	Ig kappa chain V-J
42	559	46.1	125	2 S40344	Ig kappa chain V-J
43	558	46.0	129	2 S29627	Ig kappa chain V r
44	557.5	46.0	130	2 S20637	Ig kappa chain V r
45	555.5	45.8	121	2 S40327	Ig kappa chain - h

ALIGNMENTS

RESULT 1
JE0242
Ig kappa chain NIG26 precursor - human
C:Species: Homo sapiens (man)
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0242
R:Alim, M.A.; Yamaki, S.; Hossain, M.S.; Takeda, K.; Kojima, M.; Takashi, I.; Shinoda
submitted to JIPID, November 1998
A:Description: Structure relationship of kappatype light chains with AL amyloidosis:
A:Reference number: JE0241
A:Accession: JE0242
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:16-91/Domain: immunoglobulin homology <IMM>

Query Match 84.0%; Score 1018.5; DB 2; Length 215;
Best Local Similarity 92.6%; Pred. No. 3.3e-62;
Matches 199; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY	21	EIVLTQSPATISLSPGERATLSCTASQSVS-SYLAWYQOKPGQAPRLIYDASNRATGIP 79
DB	1	EIVLTQSPATISLSPGERATLSCTASQSVSNNTLAWYQOKPGQAPSLIYDASSRATGIP 60
QY	80	PRESGSGSGTDTLTISRLPEDEVALYCCQYFTPTPGGTRLEIKRTVAAPSVFIFP 139
DB	61	DRFSGSGSGTDTLTISGLEPEDEFAVYCCQYDRPPWTGOGTKVEIKRTVAAPSVFIFP 120
QY	140	PSDEQLKSGTASVCLNLFYPREAKYQKVDNALQSGNSQESVTEQDSKDSSTLSSTL 199
DB	121	PSDEQLKSGTASVCLNLFYPREAKYQKVDNALQSGNSQESVTEQDSKDSSTLSSTL 180
QY	200	TLISKADYEKKYACEVTHQGLSSPVTKSFNRGEC 234
DB	181	TLISKADYEKKYACEVTHQGLSSPVTKSFNRGEC 215

RESULT 2
JE0244
Ig kappa chain NIG2 precursor - human
C:Species: Homo sapiens (man)
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0244
R:Alim, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazl, H.
submitted to JIPID, November 1998
A:Description: A new subgroup of k type light chains (VKV) identified in cases of AL
A:Reference number: JE0243
A:Accession: JE0244
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

Page 3

514237 chain precursor (15C5) - mouse
 ST kappa musculus (house mouse)
 10-Nov-1995 #text_change 21-Jan-2000
 D. Colten, D.

OY 185 EODSKDSTYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
 Db 181 DODSKKSTYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
 S37484

RESULT 10
 Ig kappa chain - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 R:Accession: S37484
 Submitted to the EMBL Data Library, February 1993
 A:Reference number: S37483
 A:Accession: S37484
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-225 <DUC>
 A:Cross-references: EMBL:X70424; NID:9406254; PIDN:CAA49869.1; PID:9406255
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match
 Best Local Similarity 59.9%; Score 726.5; DB 2; Length 225;
 Matches 137; Conservative 38; Mismatches 48; Indels 1; Gaps 1;

OY 11 LILWLPDTGEIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYD 70
 Db 3 LILWLPDTGEIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYD 70
 OY 71 ASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTV 62
 Db 63 ASSRYTGVPDRFTGSGYGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTV 62
 OY 131 AAPSVFTEPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKD 130
 Db 122 AAPSVFTEPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKD 121
 OY 191 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
 Db 182 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 11
 S68241

Ig kappa chain V region (Mab13-1) - mouse (fragment)
 N:Alternate names: immunoglobulin light chain
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 R:Accession: S68241; S68214
 Submitted to the EMBL Data Library, March 1994
 A:Description: Specific peroxidase activity by formation of an antibody L-chain-porphyrin
 A:Reference number: S68241
 A:Accession: S68241
 A:Molecule type: mRNA
 A:Residues: 1-218 <TAK>
 R:Cross-references: EMBL:D29670; NID:9473962; PIDN:BA06141.1; PID:9473963
 FEBS Lett. 375, 273-276, 1995
 A:Title: Thermolabile peroxidase activity with a recombinant antibody L chain-porphyrin
 A:Reference number: S68214; MUID:96085223; PMID:7498516
 A:Accession: S68214
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-218 <TAK>
 A:Cross-references: EMBL:D29670
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: immunoglobulin

Query Match
 Best Local Similarity 59.2%; Score 717; DB 2; Length 218;
 Matches 137; Conservative 31; Mismatches 46; Indels 4; Gaps 1;

OY 21 EIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYDASNRAT 76
 Db 1 EIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYDASNRAT 76
 OY 77 GIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTVAAPSVF 60
 Db 61 GIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTVAAPSVF 60
 OY 137 IFPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKSTYSL 136
 Db 121 IFPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKSTYSL 120
 OY 197 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
 Db 181 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 12
 S38865

Ig kappa chain - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Jun-2001
 R:Kipp, B.; Becker, W.; Schlaak, M.
 Submitted to the EMBL Data Library, November 1993
 A:Description: Combination of a defined specificity and desired isotype by cloning of
 A:Reference number: S38865
 A:Accession: S38865
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-219 <KIP>
 A:Cross-references: EMBL:Z27396; NID:9416538; PIDN:CA81787.1; PID:9416539
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match
 Best Local Similarity 58.8%; Score 712.5; DB 2; Length 219;
 Matches 133; Conservative 37; Mismatches 44; Indels 5; Gaps 1;

OY 21 EIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYDASNRAT 75
 Db 1 EIVLTQSPATLSLSPGERATLSCASQSVSYLAAMYQOKPGQAPRLIYDASNRAT 75
 OY 76 GIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTVAAPSV 60
 Db 61 GIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTPTGQTRLEIKRTVAAPSV 60
 OY 136 IFPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKSTYSL 135
 Db 121 IFPPSDEQLKSGTASVCLLNNFYPRKAKQWKVDNALQSGNSQESVTEQDSKSTYSL 120
 OY 196 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
 Db 181 STYSLSTLTSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234

RESULT 13
 A31790

Ig kappa chain V region (17/9) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jan-2000
 R:Accession: A31790
 J. Biol. Chem. 263, 17100-17105, 1988
 A:Title: Preliminary crystallographic data, primary sequence, and binding data for an
 A:Reference number: A92686; MUID:89034213; PMID:3182835
 A:Accession: A31790
 A:Molecule type: mRNA
 A:Residues: 1-220 <SCH>
 A:Cross-references: GB:M23626; GB:J04061; NID:9533234; PIDN:AAA39162.1; PID:9533235
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:54:35 ; Search time 13.9057 Seconds
(without alignments)
697.946 Million cell updates/sec

Title: us-09-806-276a-1

Perfect score: 1212

Sequence: 1 MEAPAQQLFLLMLPDTG.....EVTGGLSSPYTKSFNRGEC 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589.5	48.6	129	1	KV3L_HUMAN P18135 homo sapien
2	583.5	48.1	129	1	KV3M_HUMAN P18136 homo sapien
3	573	47.3	128	1	KV3K_HUMAN P06311 homo sapien
4	559.5	46.2	129	1	KV3H_HUMAN P04207 homo sapien
5	548	45.2	106	1	KAC_HUMAN P01834 homo sapien
6	545	45.0	115	1	KV3I_HUMAN P04433 homo sapien
7	490.5	40.5	116	1	KV3J_HUMAN P04434 homo sapien
8	479.5	39.6	109	1	KV3B_HUMAN P01620 homo sapien
9	477.5	39.4	109	1	KV3D_HUMAN P01622 homo sapien
10	470.5	38.8	109	1	KV3E_HUMAN P01623 homo sapien
11	456.5	37.7	108	1	KV3A_HUMAN P01619 homo sapien
12	453.5	37.4	109	1	KV3G_HUMAN P04206 homo sapien
13	444	36.6	129	1	KV1W_HUMAN P04431 homo sapien
14	442	36.5	134	1	KV4C_HUMAN P06314 homo sapien
15	439.5	36.3	109	1	KV3F_HUMAN P01624 homo sapien
16	427	35.2	131	1	KV3I_MOUSE P01661 mus musculu
17	424.5	35.0	100	1	KV3C_HUMAN P01621 homo sapien
18	423	34.9	132	1	KV3F_MOUSE P01658 mus musculu
19	422.5	34.9	133	1	KV4B_HUMAN P06313 homo sapien
20	422	34.8	114	1	KV4A_HUMAN P01625 homo sapien
21	419	34.6	117	1	KV1J_HUMAN P01602 homo sapien
22	416.5	34.4	133	1	KV2F_HUMAN P06310 homo sapien
23	410	33.8	108	1	KV1H_HUMAN P01600 homo sapien
24	401	33.1	149	1	KV5A_MOUSE P01633 mus musculu
25	400	33.0	108	1	KV1M_MOUSE P01605 homo sapien
26	399	32.9	128	1	KV5E_MOUSE P01637 mus musculu
27	399	32.9	129	1	KV1X_HUMAN P04432 homo sapien
28	397	32.8	108	1	KV1B_HUMAN P01594 homo sapien
29	396	32.7	108	1	KV1V_HUMAN P04430 homo sapien
30	395	32.6	108	1	KV1A_HUMAN P01593 homo sapien
31	395	32.6	108	1	KV1O_HUMAN P01607 homo sapien
32	393	32.4	121	1	KV4O_HUMAN P06312 homo sapien
33	392	32.3	136	1	KV5B_MOUSE P01634 mus musculu

34	389	32.1	108	1	KV1L_HUMAN P01604 homo sapien
35	388	32.0	108	1	KV1E_HUMAN P01597 homo sapien
36	388	32.0	108	1	KV1K_HUMAN P01603 homo sapien
37	388	32.0	108	1	KV1N_HUMAN P01606 homo sapien
38	388	32.0	108	1	KV1Y_HUMAN P03062 homo sapien
39	385	31.8	108	1	KV1C_HUMAN P01595 homo sapien
40	383	31.6	111	1	KV3H_MOUSE P01660 mus musculu
41	382.5	31.6	117	1	KV2E_HUMAN P06309 mus musculu
42	381.5	31.5	107	1	KV1D_HUMAN P01596 homo sapien
43	381.5	31.5	113	1	KV2D_HUMAN P01617 homo sapien
44	381	31.4	108	1	KV1Q_HUMAN P01609 homo sapien
45	381	31.4	111	1	KV3M_MOUSE P01665 mus musculu

ALIGNMENTS

RESULT 1	ID	Sequence	STANDARD	PRT	Length
KV3L_HUMAN	KV3L_HUMAN	PIR; P0022; K3HUA.			129 AA.
AC	P18135	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990	(Rel. 16, Last sequence update)			
DT	15-JUL-1999	(Rel. 38, Last annotation update)			
DE	Ig kappa chain V-III region HAH precursor.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=88171307; PubMed=3127527;				
RX	Klips T.J., Tomhave E., Chen P.P., Carson D.A.;				
RA	"Autoantibody-associated kappa light chain variable region gene				
RT	expressed in chronic lymphocytic leukemia with little or no somatic				
RT	mutation. Implications for etiology and immunotherapy."				
RT	J. Exp. Med. 167:840-852(1988).				
RL	-1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M				
CC	AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC				
CC	LEUKEMIA.				
CC	PIR; P0022; K3HUA.				
DR	HSSP; P80362; IWTL.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR003596; Ig_V.				
DR	Pfam; PF00047; Ig; 1.				
DR	SMART; SM00406; IGV; 1.				
KW	Immunoglobulin V region; Signal.				
FT	SIGNAL	1	20		
FT	CHAIN	21	129		
FT	DOMAIN	21	43		
FT	DOMAIN	44	55		
FT	DOMAIN	56	70		
FT	DOMAIN	71	77		
FT	DOMAIN	78	109		
FT	DOMAIN	110	118		
FT	DOMAIN	119	129		
FT	DISULFID	43	109		
FT	NON_TER	129	129		
FT	SEQUENCE	129 AA;	14073 MW;	D3C5529272774D0 CRC64;	
QY	Query Match	48.6%;	Score 589.5;	DB 1;	Length 129;
QY	Best Local Similarity	89.9%;	Pred. No. 2.9e-43;		
QY	Matches 116;	Conservative	5;	Mismatches 7;	Indels 1;
QY	1	MEAPAQQLFLLMLPDTG	EVLTQSPATLSLSPGERATISCRASQSV	-SSYLAWYQOK	59
QY	1	MEAPAQQLFLLMLPDTG	EVLTQSPATLSLSPGERATISCRASQSV	-SSYLAWYQOK	60
QY	60	PGQAPRLIYDASNRATGIPR	SGSGSGTDTLTISRLEPEDVALYCCQGYFTPT	PTFG	119
QY	61	PGQAPRLIYDASNRATGIPR	SGSGSGTDTLTISRLEPEDVALYCCQGYFTPT	PTFG	120
QY	120	QGTRLLEIKR	128		

```
Db 121 QGTKEIKR 129
|||||
RESULT 2
KV3M_HUMAN STANDARD; PRT; 129 AA.
AC P18136;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region HIC precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=88171307; PubMed=3127527;
RX Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;
RT "Autoantibody-associated kappa light chain variable region gene
RT expressed in chronic lymphocytic leukemia with little or no somatic
RT mutation. Implications for etiology and immunotherapy.";
RL J. Exp. Med. 167:840-852(1988).
CC -1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M
CC AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC
CC LEUKEMIA.
DR PIR; PL0021; K3HUI.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-II REGION HIC.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 56 70 FRAMEWORK-2.
FT DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JKI SEGMENT.
FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14070 MW; 7395528EA2B74D6 CRC64;
Query Match 48.1%; Score 583.5; DB 1; Length 129;
Best Local Similarity 88.4%; Pred. No. 9.5e-43;
Matches 114; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
QY 1 MEAPAQLLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQOK 59
Db 1 METPAQLFLLLMLPDTGTEIVLTQSPGTLISLSPGERATLSCRASQSVSSYLAWYQOK 60
QY 60 PGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPTFG 119
Db 61 PGQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPTFG 120
QY 120 QGTKEIKR 128
Db 121 QGTKEIKR 129
RESULT 3
KV3K_HUMAN STANDARD; PRT; 128 AA.
AC P06311;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region IARC/BL41 precursor.
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=86041852; PubMed=2997711;
RX Kiobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III.";
RL Nucleic Acids Res. 13:6499-6513(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z00021; CAA77316.1; ..
DR PIR; A01899; K3HUI.
DR HSSP; P01607; 1RET.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 128 IG KAPPA CHAIN V-II REGION IARC/BL41.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 109 117 FRAMEWORK-3.
FT DOMAIN 118 128 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 108 JKI SEGMENT.
FT NON_TER 128 128 BY SIMILARITY.
SQ SEQUENCE 128 AA; 14070 MW; CC8957F0FE3B9012 CRC64;
Query Match 47.3%; Score 573; DB 1; Length 128;
Best Local Similarity 87.5%; Pred. No. 7.3e-42;
Matches 112; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 1 MEAPAQLLFLLLMLPDTGTEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOK 60
Db 1 METPAQLFLLLMLPDTGTEIVLTQSPGTLISLSPGERATLSCRASQSVSSYLAWYQOK 60
QY 61 GQAPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPTFG 120
Db 61 GQSPRLIYDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCCQYFTPTPTFG 120
QY 121 GTRLEIKR 128
Db 121 GTRLEIKR 128
RESULT 4
KV3H_HUMAN STANDARD; PRT; 129 AA.
AC P04207;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region CLL precursor (Rheumatoid factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=86177570; PubMed=3083417;
RX Jirik F.R., Sorge J., Fong S., Heitzmann J.G., Curd J.G., Chen P.P.,
```

Wed Apr 23 13:51:10 2003

us-09-806-276a-1.1sp

Goldfien R., Carson D.A.:
"Cloning and sequence determination of a human rheumatoid factor light-chain gene."
Proc. Natl. Acad. Sci. U.S.A. 83:2195-2199(1986).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M12740; AAA58992.1;
PIR: A01898; K3HUC1.
HSSP: P80362; 1WTL.
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003596; Ig_V.
PIfam: PF00047; 1g; 1.
SMART: SM00406; 1g; 1.
Immunoglobulin V region; Signal.
KW SIGNAL 1 20
FT CHAIN 1 129 IG KAPPA CHAIN V-III REGION CBL.
FT CHAIN 21 129 FRAMEWORK-1.
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 54 FRAMEWORK-2.
FT DOMAIN 55 69 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14275 MW; 5C13B411B60CC14 CRC64;
Query Match 46.2%; Score 559.5; DB 1; Length 129; 1;
Best Local Similarity 85.3%; Pred. No. 1e-40; 10; Indels 1; Gaps 1;
Matches 110; Conservative 8; Mismatches 10;
OY 1 MEAPQQLFLLLMLPTGTEIVLQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
Db 1 MEAPQQLFLLLMLPTGTEIVLQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
OY 61 GOAPRLIYDASNRATGIPRSGSGSGTDFLTISRLEPEDVALYCCOQYFT-TPYTFG 119
Db 61 GOAPRLIYDASNRATGIPRSGSGSGTDFLTISRLEPEDVALYCCOQYFT-TPYTFG 120
OY 120 QGTRLEIKR 128
Db 121 QGTRVEIKR 129
OY 121 QGTRVEIKR 129
Db 121 QGTRVEIKR 129
RESULT 5
KAC_HUMAN STANDARD: PRT; 106 AA.
ID KAC_HUMAN
AC P01834; 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig kappa chain C region.
DE IGKC.
GN Homo sapiens (Human).
OS Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Catarrhini; Homidae; Homo.
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE (MELOMA PROTEIN EU).
RP MEDLINE=71064023; PubMed=5489770;
RX Gotlib P.D., Cunningham B.A., Rutishauser U., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VI. Amino
RL acid sequence of the light chain."
RP Biochemistry 9:3155-3161(1970).
RP [2]
RP DISULFIDE BONDS.

RX MEDLINE=71064023; PubMed=4923144;
RA Gall W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X.
RL intrachain disulfide bonds."
RN Biochemistry 9:3188-3196(1970).
[3]
RP SEQUENCE (BENCE-JONES PROTEIN TI).
RX MEDLINE=72188439; PubMed=5027703;
RT Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;
RL "Rule of antibody structure. The primary structure of Bence-Jones protein
immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein
TI). IV. The complete amino acid sequence and its significance for
the mechanism of antibody production."
RN Hoppe-Seyler's Z. Physiol. Chem. 353:189-208(1972).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=81042304; PubMed=6775818;
RT Maizel J.V., Jr., Leder P.;
RL "Cloned human and mouse kappa immunoglobulin constant and J region
genes conserve homology in functional segments."
RN Cell 22:197-207(1980).
[5]
RP SEQUENCE (BENCE-JONES PROTEIN ROY).
RX Hilschmann N., Barnikol H.U., Hess M., Langer B., Ponstingl H.,
RL Steinhilber-Kayne M., Suter L., Watanabe S.;
RN (in) Franek F., Snugar D. (eds.);
RL Gamma globulins: structure and function, pp.57-74, Academic Press,
New York (1969).
[6]
RP SEQUENCE (BENCE-JONES PROTEIN CUM).
RX MEDLINE=68242259; PubMed=5586923;
RL Hilschmann N.;
RN "The complete amino acid sequence of Bence Jones protein Cum (Kappa-
type)."
RT Hoppe-Seyler's Z. Physiol. Chem. 348:1718-1722(1967).
[7]
RP SEQUENCE (BENCE-JONES PROTEIN AG).
RX MEDLINE=69234734; PubMed=4893682;
RL Titani K., Shinoda T., Putnam F.W.;
RN "The amino acid sequence of a kappa type Bence-Jones protein. 3. The
complete sequence and the location of the disulfide bridges."
J. Biol. Chem. 244:3550-3560(1969).
[8]
RP SEQUENCE (WALDENSTROM'S MACROGLOBULIN OV).
RX MEDLINE=70201507; PubMed=5447531;
RL Kohler H., Shimizu A., Paul C., Putnam F.W.;
RN "Macroglobulin structure: variable sequence of light and heavy
chains."
Science 169:56-59(1970).
[9]
RP MISCELLANEOUS: THE ED SEQUENCE HAS THE INV (3) ALLOTYPIC MARKER,
-1- MISCELLANEOUS: THE ED SEQUENCE HAS THE INV (1,2) ALLOTYPIC
45-ALA & 83-VAL. THE ROY SEQUENCE HAS THE INV (1,2) ALLOTYPIC
MARKER, 45-ALA AND 83-LEU.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
in the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: J00241; AAA58989.1;
PIR: A02116; K3HU.
HSSP: P01842; 7FAB.
GeneW: HGNC:5716; IGKC.
MIM: 147200;
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003597; Ig_c1.
PIfam: PF00047; 1g; 1.
SMART: SM00407; 1g; 1.
PROSITE: PS00290; 1g; 1.
Immunoglobulin domain; Immunoglobulin C region.

FT NON_TER 1
FT DISULFID 26
FT DISULFID 106
FT VARIANT 83
FT VARIANT 83
FT CONFLICT 14
FT CONFLICT 57
SQ SEQUENCE 106 AA; 11609 MW; 51984D1FDD372CE8 CRC64;
INTERCHAIN (WITH A HEAVY CHAIN).
V -> L (IN INV(1,2) MARKER).
/FTID=VAR_003897.
D -> N (IN REF. 7 AND 8).
E -> Q (IN REF. 5 AND 6).

Query Match
Best Local Similarity 45.2%; Score 548; DB 1; Length 106;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 TVAPSVFIEPPSDQLKSGTASVYCLLNFPYPRKAKVQWKVDNALQSGNSQESVTEQDS 188
DB 1 TVAPSVFIEPPSDQLKSGTASVYCLLNFPYPRKAKVQWKVDNALQSGNSQESVTEQDS 188
QY 189 KDSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 234
DB 61 KDSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC 106

RESULT 6
KV3J_HUMAN

ID KV3J_HUMAN STANDARD; PRT; 115 AA.
AC P04433;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087932; PubMed=6440122;
RA Pech M., Zachau H.G.;
RT "Immunoglobulin genes of different subgroups are interdigitated
within the VK locus.";
RL Nucleic Acids Res. 12:9229-9236(1984).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X01668; NOT_ANNOTATED_CDS.
DR PIR: A01901; K3HUVG.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV_1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20
FT DOMAIN 21
FT DOMAIN 21
FT DOMAIN 21
FT DOMAIN 44
FT DOMAIN 54
FT DOMAIN 55
FT DOMAIN 69
FT DOMAIN 70
FT DOMAIN 76
FT DOMAIN 108
FT DISULFID 109
FT DISULFID 115
FT NON_TER 115
SQ SEQUENCE 115 AA; 12575 MW; 2DE47CDA3A17D555 CRC64;

Query Match
Best Local Similarity 45.0%; Score 545; DB 1; Length 115;
Matches 107; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 60
DB 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 60
QY 61 GOAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115
DB 61 GOAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115

RESULT 7
KV3J_HUMAN

ID KV3J_HUMAN STANDARD; PRT; 116 AA.
AC P04434;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087932; PubMed=6440122;
RA Pech M., Zachau H.G.;
RT "Immunoglobulin genes of different subgroups are interdigitated
within the VK locus.";
RL Nucleic Acids Res. 12:9229-9236(1984).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X02725; NOT_ANNOTATED_CDS.
DR PIR: A01901; K3HUVG.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV_1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20
FT DOMAIN 21
FT DOMAIN 21
FT DOMAIN 21
FT DOMAIN 44
FT DOMAIN 55
FT DOMAIN 56
FT DOMAIN 70
FT DOMAIN 71
FT DOMAIN 77
FT DOMAIN 109
FT DISULFID 110
FT DISULFID 116
FT NON_TER 116
SQ SEQUENCE 116 AA; 12575 MW; 51CD55BA53B21929 CRC64;

Query Match
Best Local Similarity 40.5%; Score 490.5; DB 1; Length 116;
Matches 98; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

QY 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 59
DB 1 MEAPQQLFLLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQOKP 59
QY 60 PGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115
DB 61 PGQAPRLIYDASNRATGIPRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTP 115

RESULT 8
KV3B_HUMAN

RN		[1]	SEQUENCE.	
RP			MEDLINE=82046598; pubmed=6794615;	
RX			Andrews D.W., Capra J.D.: MEDIANE=82046598; pubmed=6794615;	
RA			"Amino acid sequence of the variable regions of light chains from two	
RT			idiotypically cross-reactive human IgM anti-gamma-globulins of the wa	
RT			group.";	
RT			Biochemistry 20:5816-5822(1981).	
RL			-I MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA	
CC			GLOBULIN ACTIVITY.	
CC			PIR; A01892; K3HDSI.	
DR			HSSP; P80362; 1WTL.	
DR			InterPro; IPR003006; IG_MHC.	
DR			InterPro; IPR003596; IG_V.	
DR			Pfam; PF00047; IG_1.	
DR			SMART; SM00406; IGV_1.	
DR			Immunoglobulin V region.	
KW		23	BY SIMILARITY.	
KW		DISULFID		
FT		NON TER		
FT		SEQUENCE		
SO				

RESULT 9	STANDARD:	PRT:	109 AA.
KV3D_HUMAN			
ID	P01622;		
AC	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, last sequence update)		
DT	21-JUL-1986 (Rel. 38, last annotation update)		
DT	15-JUL-1999 (Rel. 38, last annotation update)		
DT	15-JUL-1999 (Rel. 38, last annotation update)		
DE	Ig kappa chain V-III region TI.		
OS	Homo sapiens (Human).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=72188439; PubMed=5027703;		
RA	Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;		
RT	"Rule of antibody structure. The primary structure of a monoclonal		
RT	immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein		
RT	Ti). IV. The complete amino acid sequence and its significance for		
RT	the mechanism of antibody production.";		
RT	Hoppe-Seyler's Z. Physiol. Chem. 353:189-208(1972).		
RL	THE MECHANISM OF ANTIBODY PRODUCTION.		
CC	-1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.		
CC	-1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.		
DR	PIR: A01895; K3HUT1.		
DR	HSSP: P80362; 1WTL.		
DR	InterPro: IPR003006; Ig_MHC.		
DR	InterPro: IPR003596; Ig_V.		
DR	Pfam: PF00047; Ig_1.		
DR	SMART: SM00406; IGV; 1.		
DR	SMART: SM00406; IGV; 1.		

	Query Match	Similarity	86.2%	8;	Mismatches	6;	100
	Best	Local	Conservative				
	Matches	94;					
QY	21	EIVLTOSPATLSLSPGERATLSCRASQSVS-SYLAWYQOKPGQAPRLLIYDASNRATGIP	79				
		EIVLTOSPATLSLSPGERATLSCRASQSVSNSFLAWYQOKPGQAPRLLIYVASSRATGIP	60				
Db	1	EIVLTOSPATLSLSPGERATLSCRASQSVSNSFLAWYQOKPGQAPRLLIYDASNRATGIP	128				
QY	80	PRFSGSGSGTDFTLTISRLEPEDVALYYCOOYFTPTPYFGOGTRLEIKR	109				
Db	61	DRFSGSGSGTDFTLTISRLEPEDFAVYYCOOYGSSEPSTFGOGTKVELKR					

RESULT	10			
KV3E_HUMAN	STANDARD;	PRT;	109	AA.
ID	P01623;			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, last sequence update)			
DT	15-JUL-1999 (Rel. 38, last annotation update)			
DT	15-JUL-1999 (Rel. 38, last annotation update)			
DE	Ig kappa chain V-III region WOL.			
OS	Homo sapiens (Human); Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.			
OC	Mammalia; Eutheria;			
NCBI	TaxID=9606;			

SEQUENCE: Pubmed-67994615;
MEDLINE-82046598; Capra J.D.;
Andrews D.W.; Capra J.D.;
Amino acid sequence of the variable regions of the wa
idiotypically cross-reactive human Igm anti-gamma-globulins of the wa
group.
Biochemistry 20:5816-5822(1981).
-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
GLOBULIN ACTIVITY.
PIR: A01896; K3HUML.
DR PIR: A01896; K3HUML.
DR HSSP: P80362; 1MTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR InterPro: IPR00047; Ig; 1.
DR Pfam: PF000406; IGV; 1.
DR SMART: SM00406; V region.
DR Immunoglobulin V region.
BY SIMILARITY.
KW DISULEID 23 89
FT 109 109
FT NON_TER 109 AA; 11746 MW; 566C115E6B9CBECE CRC64;
SQ SEQUENCE 109 AA; 11746 MW; 566C115E6B9CBECE CRC64;
38.8%; Score 470.5; DB 1; Length 109;
No 2.8e-33; 1; Gaps 1;

Query Match	Similarity	86.2%	Pred. NO. 9	Indels
Best Local	Conservative	5	Mismatches	9
Matches	94			
OY	21	EIVLTQSPATLSLSPGERATLSCRASQSVSS-YLAWYQOKPGQAPRLLIYDASNRAWGIP	128	79
		EIVLTQSPGLSLSPGERATLSCRASQSVSSGILGWYQOKPGQAPRLLIYGASSRATGIP	128	60
OY	80	PRFSGSGGTDFLTITSLRLEPEDVALYCCQYFTTPYTFGGGTRLEIKR	109	
Db	61	DRFSGSGGTDFLTITSLRLEPEDFAYVYCCQYSLGRFTGGGTRLEIKR	109	
RESULT	11			
KV3A_HUMAN		STANDARD:	PRT:	108 AA.
ID	KV3A_HUMAN			
AC	P01619			
DT	21-JUL-1986 (Rel. 01, last sequence update)			
DT	21-JUL-1986 (Rel. 01, last annotation update)			
DT	15-JUL-1999 (Rel. 38, last annotation update)			
DE	19 kappa chain V-III region B6.			
OS	Homo sapiens (Human).			

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE.
RA Milstein C.;
RT "The basic sequences of immunoglobulin kappa chains: sequence studies
of Bence Jones proteins Rad, Fr4 and B6.";
RL FEBS Lett. 2:301-304(1969).
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR; A01891; K3HUB6.
DR HSSP; P80362; 1WTU.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; 1g; 1.
DR SMART; SM00406; 1g; 1.
KW Immunoglobulin V region.
FT DISULFID 23 89
FT NON_TER 108 108
SO SEQUENCE 108 AA; 11635 MW; 8BC1AF07A419E3D CRC64;

Query Match
Best Local Similarity 37.7%; Score 456.5; DB 1; Length 108;
Matches 85; Conservative 14; Mismatches 8; Indels 1; Gaps 1;

QY 21 EIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKPGQAPRLITYDASNRATGIP 79
DB 1 ZIVLTSPGTLSPGZRALISCRASQSLSGNYLAWYQKPGQAPRLITYDASNRATGIP 60
QY 80 PRFSGSGSTDEFTLTISRLPEPDVALYCCQYFTPTPTGQGTREIK 127
DB 61 DRFSGSGSGADFTLTISRLPEPDVALYCCQYFTPTPTGQGTREIK 108

RESULT 12
KV3G_HUMAN
ID KV3G_HUMAN
AC P04206; STANDARD; PRT; 109 AA.

DT 20-MAR-1987 (Rel. 04, Created)
DE 15-JUL-1999 (Rel. 04, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE.
RA MEDLINE-86230578; PubMed-3086710;
RT "Amino acid sequence of a light chain variable region of a human
rheumatoid factor of the Wa idiotype group, in part predicted by its
Mol. Immunol. 23:239-244(1986).
DR PIR; A01893; K3HUB6.
DR HSSP; P80362; 1WTU.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; 1g; 1.
DR SMART; SM00406; 1g; 1.
KW Immunoglobulin V region.
FT DISULFID 23 89
FT NON_TER 109 109
SO SEQUENCE 109 AA; 11830 MW; 9349A5B1D93588B6 CRC64;

Query Match
Best Local Similarity 37.4%; Score 453.5; DB 1; Length 109;
Matches 89; Conservative 9; Mismatches 10; Indels 1; Gaps 1;
QY 21 EIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKPGQAPRLITYDASNRATGIP 79
DB 1 EIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKPGQAPRLITYDASNRATGIP 60

QY 80 PRFSGSGSTDEFTLTISRLPEPDVALYCCQYFTPTPTGQGTREIK 128
DB 61 DRFSGSGSGTDEFTLTISRLPEPDVALYCCQYFTPTPTGQGTREIK 109

RESULT 13
KV1W_HUMAN
ID KV1W_HUMAN
AC P04431; STANDARD; PRT; 129 AA.

DT 13-AUG-1987 (Rel. 05, Created)
DE 15-JUL-1999 (Rel. 05, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RA MEDLINE-85014148; PubMed-6091049;
RT "Immunoglobulin genes of the kappa light chain type from two human
lymphoid cell lines are closely related.";
RL Nucleic Acids Res. 12:6995-7006(1984).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X00965; CAA25477.1; ALT_TER.
PIR; A01883; K1HUMK.
DR HSSP; P01607; 1REI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; 1g; 1.
DR SMART; SM00406; 1g; 1.
KW Immunoglobulin V region; Signal.
FT CHAIN 1 22
FT DOMAIN 23 129
FT DOMAIN 23 129
FT DOMAIN 46 45
FT DOMAIN 57 56
FT DOMAIN 72 71
FT DOMAIN 79 78
FT DOMAIN 111 110
FT DOMAIN 120 119
FT DISULFID 45 129
FT NON_TER 129 129
SO SEQUENCE 129 AA; 14069 MW; F941FA07D4AFC2F9 CRC64;

Query Match
Best Local Similarity 36.6%; Score 444; DB 1; Length 129;
Matches 86; Conservative 17; Mismatches 24; Indels 0; Gaps 0;
QY 1 MEAPQLFLILMLPDTGIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQKPG 60
DB 3 MRVPAQLGLLILMLKARCDIOMTQSSLSASVGDVYITTCRASQSSISNYLAWYQKPG 62
QY 61 GQAPRLITYDASNRATGIPPRFSGSGSTDEFTLTISRLPEPDVALYCCQYFTPTPTGQ 60
DB 63 GKAPRLITYDASNRATGIPPRFSGSGSTDEFTLTISRLPEPDVALYCCQYFTPTPTGQ 62
QY 121 GTRLEIK 127
DB 123 GTRLEIK 129

RESULT 14
KV4C_HUMAN

DT		21-JUL-1986.	(Rel. 01, last sequence update)
DT		16-OCT-2001	(Rel. 40, last annotation update)
DE		Ig kappa chain V-II	region POM.
OS		Homo sapiens (Human).	
OC		Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;
OC		Mammalia; Eutheria;	Primates; Catarrhini; Homnidae; Homo.
OX		NCBI_TaxID=9606;	
RN	[1]		
RP		SEQUENCE.	
RX		MEDLINE=76276460;	PubMed=60899;
RA		Klapper D.G., Capra J.D.;	
RT		"The amino acid sequence of the variable regions of the light chains	
RL		from two idiotypically cross reactive Igm anti-gamma globulins."	
CC		Ann. Immunol. (Paris) 127C:261-271(1976).	
CC	-1-	MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA	
CC		GLOBULIN ACTIVITY.	
DR	PIR;	A01897; K3HUPM.	
DR	HSSP;	P80362; 1WTL.	
DR	InterPro;	IPR003006; Ig_MHC.	
DR	InterPro;	IPR003596; Ig_V.	
DR	Pfam;	PF00047; Ig_1.	
DR	SMART;	SM00406; IGv_1.	
KW	Immunoglobulin V region.		
FT	DISULFID	23 89	BY SIMILARITY.
FT	NON_TER	109 109	
SQ	SEQUENCE	109 AA; 11922 MW; 62821DDC6A8ABA86 CRC64;	

Query Match	36.3%;	Score 439.5;	DB 1;	Length 109;
Best Local Similarity	79.8%;	Pred. No. 1.2e-30;		
Matches	87;	Conservative	8;	Mismatches 13;
				Indels 1;
				Gaps 1.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:57:05 ; Search time 25.8934 Seconds
(without alignments)
1862.055 Million cell updates/sec

Title: US-09-806-276A-1

Perfect score: 1212
Sequence: 1 MEAPAQQLFLLLWLPDPTG.....EVTTHQGLSSPVTKSFNRGEC 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	961.5	79.3	239	4 Q8TCD0	Q8TCD0 homo sapien
2	752	62.0	238	11 Q99M37	Q99M37 mus musculu
3	737	60.8	238	11 Q8VC16	Q8VC16 mus musculu
4	732	60.4	234	11 Q8VCP0	Q8VCP0 mus musculu
5	728	60.1	234	11 Q8R062	Q8R062 mus musculu
6	721	59.5	234	11 Q91WF8	Q91WF8 mus musculu
7	719.5	59.4	239	11 Q8VC35	Q8VC35 mus musculu
8	703	58.0	233	11 Q91WS9	Q91WS9 mus musculu
9	699	57.7	214	11 Q91IAS	Q91IAS mus musculu
10	698.5	57.6	235	11 Q91W12	Q91W12 mus musculu
11	665	54.9	234	11 Q8R028	Q8R028 mus musculu
12	479.5	39.6	109	4 Q9UL78	Q9UL78 homo sapien
13	457.5	37.7	233	4 Q8TBC9	Q8TBC9 homo sapien
14	454	37.5	237	4 Q8WTU6	Q8WTU6 homo sapien
15	453.5	37.4	109	4 Q9UL86	Q9UL86 homo sapien
16	452	37.3	237	4 Q8WUK4	Q8WUK4 homo sapien

17	446	36.8	108	4 Q9UL83	Q9UL83 homo sapien
18	432.5	35.7	109	4 Q9UL85	Q9UL85 homo sapien
19	410.5	33.9	236	4 Q96E61	Q96E61 homo sapien
20	407	33.6	108	4 Q9UL79	Q9UL79 homo sapien
21	405	33.4	108	4 Q9UL77	Q9UL77 homo sapien
22	398	32.8	108	4 Q9UL70	Q9UL70 homo sapien
23	397	32.8	233	4 Q96I69	Q96I69 homo sapien
24	396	32.7	235	11 Q99M11	Q99M11 mus musculu
25	395	32.6	116	4 Q96PF6	Q96PF6 homo sapien
26	390.5	32.2	107	4 Q96SA9	Q96SA9 homo sapien
27	389	32.1	240	4 Q8WUK3	Q8WUK3 homo sapien
28	380	31.4	298	11 Q9QYF0	Q9QYF0 mus musculu
29	378	31.2	111	11 Q920E9	Q920E9 mus musculu
30	375.5	31.0	106	5 Q9U410	Q9U410 schistosoma
31	375.5	31.0	134	11 Q8VDD0	Q8VDD0 mus musculu
32	366.5	30.2	107	4 Q9UL81	Q9UL81 homo sapien
33	361	29.8	109	11 Q920E6	Q920E6 mus musculu
34	360	29.7	108	11 Q8VIJ0	Q8VIJ0 mus musculu
35	355	29.3	107	11 Q9ER29	Q9ER29 mus musculu
36	353	29.1	127	11 Q925S9	Q925S9 mus musculu
37	347	28.6	99	11 Q9JL74	Q9JL74 mus musculu
38	341	28.1	114	4 Q9UL80	Q9UL80 homo sapien
39	332	27.4	103	11 Q9JL80	Q9JL80 mus musculu
40	329.5	27.2	104	11 Q9JL82	Q9JL82 mus musculu
41	329	27.1	101	11 Q9JL78	Q9JL78 mus musculu
42	327	27.0	107	11 Q9JL84	Q9JL84 mus musculu
43	322	26.6	97	11 Q9JL76	Q9JL76 mus musculu
44	303.5	25.0	241	11 Q92IA6	Q92IA6 mus musculu
45	284	23.4	109	6 Q9NOW5	Q9NOW5 oryctolagus

ALIGNMENTS

RESULT 1
Q8TCD0 PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC022362; AAH22362.1; -
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26234 MW; FAGEDC3A3B03871D CRC64;

Query Match 79.3%; Score 961.5; DB 4; Length 239;

Best Local Similarity 76.6%; Pred. No. 5.8e-83;
Matches 183; Conservative 28; Mismatches 23; Indels 5; Gaps 1;

QY	1	MEAPAQQLFLLLWLPDPTGELVLTQSPATLSLSPGERATLSCRASQSV-----SSYLAW 55
DB	1	MRLPAQLGLMLWVPGSSGDVMTQSPSLPVTIGQPASISCRSTOSLVSDGNTYLNW 60
QY	56	YQKPGQAPRLITDASNRATGIPRFSGSGSGTDFTLTISRLEPEDVALYCOQYFTTP 115
DB	61	FOORPGQSPRLITYKSNRDSGVPRFSGSGSGTDFTLITRVEAEADVGVFCMQGTHWP 120
QY	116	YTFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLNNFYPREAKVQWKYDNLQ 175
DB	121	STFGGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLNNFYPREAKVQWKYDNLQ 180
QY	176	SGNSQSVTEQDSKSTYSLSTLTLSKADYEKKHYACEVTHQGLSSPVTKSFNRGEC 234

61 GOAPRLTYDASNRTGIPRFGS

61 GQAPRLLYDASNRATGIPPRFSGSGSGTDETIITSPRI FREDVAIVVCCOQVEMEDVEMDCC 100

```

Db 61 GKSPQLLVNNAKTLADGVSRFSGSRSGTQFSLKINSLOPEDFGSYCQHHSGIPFTFGS 120
QY 121 GTRLEIKRTVAAPSVFIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
Db 121 GTRLEIKRADAPVSIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
QY 181 ESYTEQDSKDSSTYSLSSTLTLSKADYEKKHYACEVTHQGLSPVTKSFNRGEC 234
Db 181 NSWTDQDSKDSSTYSLSSTLTLTKEDEYERHNSYTCEATHKTSPIVKSFNRENC 234

RESULT 5
QY 08R062 PRELIMINARY; PRT; 234 AA.
AC 08R062;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 25.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027418; AAH27418.1;
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25857 MW; 4EB08C81426AEAB1 CRC64;

Query Match
Best Local Similarity 59.4%; Score 728; DB 11; Length 234;
Matches 139; Conservative 34; Mismatches 61; Indels 0; Gaps 0;

QY 1 MEAPAQLLFLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
Db 1 MMSAQFLGLLLCFQGTCDIQMTQTSSLSASLSDRVITSCASQGISNYLAWYQOKP 60
QY 61 GQAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTFGQ 120
Db 61 DGTAKLLIYTSLSHSGVPSRFGSGSGTHTSLTISNLEPEDATYCCQYFTPTFGS 120
QY 121 GTRLEIKRTVAAPSVFIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
Db 121 GTRLEIKRADAPVSIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
QY 181 ESYTEQDSKDSSTYSLSSTLTLSKADYEKKHYACEVTHQGLSPVTKSFNRGEC 234
Db 181 NSWTDQDSKDSSTYSLSSTLTLTKEDEYERHNSYTCEATHKTSPIVKSFNRENC 234

RESULT 6
QY 091WF8 PRELIMINARY; PRT; 234 AA.
AC 091WF8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 25.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015292; AAH15292.1;
KW InterPro; IPR003006; Ig_MHC.
```

```

DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF00047; Ig_2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
SQ SEQUENCE 234 AA; 25929 MW; B0D0B0E6EB7812D2 CRC64;

Query Match
Best Local Similarity 59.5%; Score 721; DB 11; Length 234;
Matches 138; Conservative 34; Mismatches 62; Indels 0; Gaps 0;

QY 1 MEAPAQLLFLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 60
Db 1 MMSAQFLGLLLCFQGTCDIQMTQTSSLSASLSDRVITSCASQDISNYLAWYQOKP 60
QY 61 GQAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPTFGQ 120
Db 61 DGTAKLLIYTSLSHSGVPSRFGSGSGTHTSLTISNLEPEDATYCCQYFTPTFGS 120
QY 121 GTRLEIKRTVAAPSVFIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
Db 121 GTRLEIKRADAPVSIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQSGNSQ 180
QY 181 ESYTEQDSKDSSTYSLSSTLTLSKADYEKKHYACEVTHQGLSPVTKSFNRGEC 234
Db 181 NSWTDQDSKDSSTYSLSSTLTLTKEDEYERHNSYTCEATHKTSPIVKSFNRENC 234

RESULT 7
QY 08VC55 PRELIMINARY; PRT; 239 AA.
AC 08VC55;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 26.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021781; AAH21781.1;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00409; Ig_2.
DR SMART; SM00407; IG_c1.
DR SMART; SM00406; IG_v.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26303 MW; C16119CACA25C337 CRC64;

Query Match
Best Local Similarity 59.4%; Score 719.5; DB 11; Length 239;
Matches 136; Conservative 43; Mismatches 55; Indels 5; Gaps 1;

QY 1 MEAPAQLLFLLMLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSYLAWYQOKP 55
Db 1 MKLEPVLLVLLFTSPASSDYLVTPLSLPVGNDQASISCKSTKSLNSDGFYLDW 60
QY 56 YQKRGQAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPT 115
Db 61 YLQKRGQAPRLIYDASNRATGIPRFGSGSGTDFLTISRLEPEDVALYCCQYFTPT 120
QY 116 YTFGGGTRLEIKRTVAAPSVFIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQ 175
Db 121 YTFGGGTRLEIKRTVAAPSVFIPTSPDEQLKSGTASVCLLNFFPREAKVQWKVDNALQ 180
```

QY 176 SCNSQSEVTEQDSKSTYSLSTLTLKADYEKHKVYACEVTHQGLSSPVTKSFNREGC 234
Db 181 ONGVLNSWTDQDSKSDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC 239

RESULT 8

Q91WS9 PRELIMINARY; PRT; 233 AA.
AC Q91WS9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Hypothetical 25.8 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013496; AAH13496.1;
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 233 AA; 25781 MW; B1C184DA149A16EB CRC64;

Query Match 58.0%; Score 703; DB 11; Length 233;
Best Local Similarity 58.9%; Pred. No. 1.6e-58;
Matches 132; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

QY 11 LLLWLPDTGEIVLQSPATLSLSPGERATLSGRASQSVSSYLAWYQKPGAPRLIYD 70
Db 10 LLLCFQGRCDIQMTQTSSLSASLSDRVYISCSGSGIANYLNWYQKPDGTVKLLIY 69
QY 71 ASNRATGIPRFGSGSGTDTLTISRLEPEDVALYCCQYFTPTFGGTRLEIKRTV 130
Db 70 TSSLHSGVPSRFGSGSGTDTLTISRLEPEDVALYCCQYFTPTFGGTRLEIKRAD 129
QY 131 AAPSVFIPTSPDEQLKSGTASVCLLNFPYREKAVQWKNALQSGNSQSEVTEQDSKD 190
Db 130 AAPTVSIFPPSSEQLTSGGASVCFLLNFPYKIDINVKWKIDGSERQNGVLNSWTDQDSKD 189
QY 191 STYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNREGC 234
Db 190 STYMSSTLTITKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC 233

RESULT 9

Q9RIA5 PRELIMINARY; PRT; 214 AA.
AC Q9RIA5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Kappa light chain of Mab7 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152371; AAD40242.1;
DR HSSP; P01679; 2FBJ.

DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00406; IgV_1.
DR SMART; SM00410; Ig_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 214 AA; 23922 MW; 52BA205FDE95E2A CRC64;

Query Match 57.7%; Score 699; DB 11; Length 214;
Best Local Similarity 59.3%; Pred. No. 3.4e-58;
Matches 127; Conservative 37; Mismatches 50; Indels 0; Gaps 0;

QY 21 EIVLTQSPATLSLSPGERATLSGRASQSVSSYLAWYQKPGAPRLIYDASNRATGIP 80
Db 1 DIQLTQSPSSMYASLGERVTTTCASODINSYLSWFQKPGKSPKTLIRANRLVDGPS 60
QY 81 RFGSGSGTDTLTISRLEPEDVALYCCQYFTPTFGGTRLEIKRTVAAPSVFIPT 140
Db 61 RFGSGSGGDTSLTISLEYEDMGIIYCIQYDEPFTFGSGTRLEIKRADAAPTVSIFP 120
QY 141 SDEQLKSGTASVCLLNFPYREKAVQWKNALQSGNSQSEVTEQDSKSTYSLSTLT 200
Db 121 SSEQLTSGGASVCFLLNFPYKIDINVKWKIDGSERQNGVLNSWTDQDSKSTYSMSSTLT 180
QY 201 LSKADYEKHKVYACEVTHQGLSSPVTKSFNREGC 234
Db 181 LTKDEYERHNSYTCEATHKTSTSPIVKCFNRNEC 214

RESULT 10

Q91W12 PRELIMINARY; PRT; 235 AA.
AC Q91W12;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Unknown (protein for MGC:6582).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006643; AAH06643.1;
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 235 AA; 26021 MW; 5FC73BDEBD5E8FEF CRC64;

Query Match 57.6%; Score 698.5; DB 11; Length 235;
Best Local Similarity 60.9%; Pred. No. 4.4e-58;
Matches 131; Conservative 31; Mismatches 52; Indels 1; Gaps 1;

QY 20 GEIVLTQSPATLSLSPGERATLSGRASQSVSSYLAWYQKPGAPRLIYDASNRATGIP 79
Db 22 GQIVLTQSPALMSASPGERVMTTCSASSV-SHMHWYQKSGTSPKWIYDTFKLTSGVP 80
QY 80 PRFGSGSGTDTLTISRLEPEDVALYCCQYFTPTFGGTRLEIKRTVAAPSVFIPT 139
Db 81 PRFGSGSGTSLTISNMEADVATYYCCQWSRNPPTFGVGTKLEIKRADAAPTVSIFP 140
QY 140 PSDEQLKSGTASVCLLNFPYREKAVQWKNALQSGNSQSEVTEQDSKSTYSLSTLT 199
Db 141 PSSEQLTSGGASVCFLLNFPYKIDINVKWKIDGSERQNGVLNSWTDQDSKSTYSMSSTLT 200
QY 200 TLSKADYEKHKVYACEVTHQGLSSPVTKSFNREGC 234

Db 201 TLTKEDEYERHNSYCEATHKSTSPIVKSFNNEC 235

RESULT 11

Q8R028 PRELIMINARY; PRT; 234 AA.
AC Q8R028;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028540; AAH28540.1; -
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25702 MW; 148377F9C1CD0AEE CRC64;

Query Match 54.9%; Score 665; DB 11; Length 234;
Best Local Similarity 55.2%; Pred. No. 6.4e-55;
Matches 127; Conservative 32; Mismatches 71; Indels 0; Gaps 0;
QY 5 AQLLELLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQKPGQAP 64
Db 5 APLSLLLLCVSDSRRAETTVTQSPASLSVAATGEKVTIRCTSTIDDDMNWYQKPGQAP 64
QY 65 RLIIYDASNRATGIPPRFSGSGGTDFLTISRLEPEDVALYYCQYFTPTFGGQTRL 124
Db 65 KLLISEGNTLRPGVPSRFSSSGYGTDFVFTIENTLSEADVADYCLQSDNMLTFGAGTKL 124
QY 125 EIKRTVAAPSVFIFPPSDEQLKSGTASVCLINFEYPREAKYQWKVDNALQSGNSQESVT 184
Db 125 ELKRDAAPIYSIFPPSSEQLTSGASVCFLINFEYPRKIDINVKWKIDGSEKQNGVLSWT 184
QY 185 EQDSKSTYSLSSTLTLSKADYEKHKYACEVTHQGLSPVTKSFNRGEC 234
Db 185 DQDSKSTYSMSSTLTLTKEDEYERHNSYCEATHKSTSPIVKSFNNEC 234

RESULT 12

Q9UL78 PRELIMINARY; PRT; 109 AA.
AC Q9UL78;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
OS Myosin-reactive immunoglobulin light chain variable region (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RT Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035036; AAD56272.1; -
DR HSSP; P80362; 1WTI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1 1

FT NON_TER 109 109
SQ SEQUENCE 109 AA; 11646 MW; 5F675C52EC7EE197 CRC64;

Query Match 39.6%; Score 479.5; DB 4; Length 109;
Best Local Similarity 87.2%; Pred. No. 8e-38;
Matches 95; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 21 EIVLTQSPATLSLSPGERATLSCRASQSV-SSYLAWYQKPGQAPRLIYDASNRATGIP 79
Db 1 EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQKPGQAPRLIYDASNRATGIP 60
QY 80 PRFSGSGGTDFLTISRLEPEDVALYYCQYFTPTFGGQTRLTKR 128
Db 61 DRFSGSGGTDFLTISRLEPEDCAVYYCQYQGSPLTFGGGTKEIKR 109

RESULT 13

Q8TBC9 PRELIMINARY; PRT; 233 AA.
AC Q8TBC9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022823; AAH22823.1; -
KW Hypothetical protein.
SQ SEQUENCE 233 AA; 24867 MW; 367411BFD6F4DF92 CRC64;

Query Match 37.7%; Score 457.5; DB 4; Length 233;
Best Local Similarity 43.5%; Pred. No. 2.8e-35;
Matches 101; Conservative 43; Mismatches 79; Indels 9; Gaps 6;

QY 7 LLEFLLLMLPDTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQKPGQAPRL 66
Db 6 LLEPLLLTCTGSEASYELTQ-PPSVSVSPGQYARITCSGDALPKQYAWYQKPGQAPVL 64
QY 67 LIYDASNRATGIPPRFSGSGGTDFLTISRLEPEDVALYYCQYFT--TPYTFGGQTRL 124
Db 65 VIYKDNERPSGIPERFSGSSGTVTTLTISGVQAEADYDCCOSADSSCTYVWFGGQTRL 124
QY 125 EI-KRTVAAPSVFIFPPSDEQLKSGTASVCLINFEYPREAKYQWKVDNA-LQSGNSQES 182
Db 125 TVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAG--VET 182
QY 183 VTEQDSKSTYSLSSTLTLSKADYEKHKYACEVTHQGLSPVTKSFNRGEC 234
Db 183 TTPSKQSNKKYAASSYLSLTPEQWKS HKYSQVTHG--STVEKTVAPTEC 232

RESULT 14

Q8WTU6 PRELIMINARY; PRT; 237 AA.
AC Q8WTU6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;

RA Strausberg R.; Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC022098; AAH22098.1; --
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig_2.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00407; IGC1; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 237 AA; 24884 MW; E6CF371E753968E8 CRC64;

Query Match	37.5%;	Score 454;	DB 4;	Length 237;
Best Local Similarity	43.6%;	Pred. No. 6.1e-35;		
Matches 103;	Conservative 44;	Mismatches 73;	Indels 16;	Gaps 8;

[illegible]

RESULT 15

ID	Q9UL86	PRELIMINARY;	PRT;	109 AA.
AC	Q9UL86;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Myosin-reactive immunoglobulin kappa chain variable region (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98277139; PubMed=9614934;			
RA	Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,			
RA	Young D.C.;			
RT	"Myosin-reactive autoantibodies in rheumatic carditis and normal			
RT	fetus.";			
RL	Clin. Immunol. Immunopathol. 87:184-192(1998).			
DR	EMBL; AF035028; AAD56264.1; -.			
DR	HSSP; P80362; 1WTL.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF00047; Ig; 1.			
DR	SMART; SM00406; IGV; 1.			
FT	NON_TER	1		
FT	NON_TER	109	109	
SEQUENCE	109 AA;	11928 MW;	243325F72C7DAC83	CRC64;

Query Match	37.48;	Score 453.5;	DB 4;	Length 109;
Best Local Similarity	82.68;	Pred. No. 2.3e-35;		
Matches 90; Conservative	8;	Mismatches 10;	Indels 1;	Gaps 1;

```

QY      21 EIVLQSPATLSLSPGERATLSCRASQSV-SSYLAWYQQKPGQABRLIIYDASNRAITGP 79
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1 EIVLTQSPGTLSLFGERATLSCRASQSVSSSYLAWYQOKPGQADRLIIYGTSSRAITGP 60

Qy 80 PRFGSGSGTDFTLTISRLEPEDVALYCCQYFTTPTFGQTRLEIKR 128
||||| ||||||| : :||| ||: |||
61 DRFGSGSEITDFTLTISRLEPEDFAVYYCCQYSGSITFTFGPTKVDIKR 109

Db

Search completed: April 23, 2003, 11:02:42
Job time : 27.8934 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:53:45 ; Search time 54.6516 Seconds
(without alignments)
619.298 Million cell updates/sec

```
Title: US-09-806-276A-2
Perfect score: 1326
Sequence: 1 MPAAKPIYLKANNKKGKPF.....QLDGPILDEVLNMDKNK 254
```

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1326	100.0	254	21	AAV92240	Human bone marrow-
2	562.5	42.4	538	22	ABG25106	Novel human diagn
3	256.5	19.3	137	22	ABB27392	Human peptide #43
4	256.5	19.3	137	22	ABB32540	Peptide #46 encode
5	256.5	19.3	137	22	AAW53373	Human brain expres
6	256.5	19.3	137	22	AAW13611	Peptide #45 encode
7	256.5	19.3	137	22	AAW26008	Peptide #45 encode
8	256.5	19.3	137	22	AAW01361	Peptide #43 encode
9	256.5	19.3	137	23	ABG35381	Human peptide encc
10	223	16.8	349	22	AAE01862	Mouse zmsel protei

11	222.5	16.8	356	22	AAM93984	Human stomach cancer
12	222.5	16.8	356	22	AAE01861	Human zmsel protei
13	222.5	16.8	356	22	AAB95088	Human protein sequ
14	130.5	9.8	1274	22	ABB60313	Drosophila melanog
15	130.5	9.8	1356	22	ABB67291	Drosophila melanog
16	111	8.4	149	22	AAG73979	Human colon cancer
17	109.5	8.3	1475	22	AAAM40232	Human polypeptide
18	108	8.1	1331	22	AAM39048	Human polypeptide
19	105	7.9	1283	21	AAB07205	Human Factor VIIIC
20	105	7.9	1284	6	AAP50108	Factor-VIIIC deriv
21	105	7.9	2098	17	AAR86863	Factor-VIII. Homoc
22	105	7.9	2332	8	AAP71726	Factor VIII:c varia
23	105	7.9	2332	8	AAP71727	Factor VIII:c varia
24	105	7.9	2332	8	AAP71728	Factor VIII:c varia
25	105	7.9	2332	8	AAP71729	Factor VIII:c varia
26	105	7.9	2332	14	AAR43257	Human Factor VIII
27	105	7.9	2332	18	AAM33222	Procoagulant-activ
28	105	7.9	2332	18	AAM33223	Procoagulant-activ
29	105	7.9	2332	18	AAM33224	Procoagulant-activ
30	105	7.9	2332	18	AAM33225	Procoagulant-activ
31	105	7.9	2332	18	AAM33226	Procoagulant-activ
32	105	7.9	2332	19	AAM53483	Human factor VIII
33	105	7.9	2332	19	AAM44132	Homo sapiens modif
34	105	7.9	2332	20	AAY31594	Human factor VIII
35	105	7.9	2332	21	AAY57847	Human Factor VIII
36	105	7.9	2332	22	AAE10826	Human mature wild
37	105	7.9	2332	22	AAE11200	Human factor VIII
38	105	7.9	2332	22	AAB71902	N-terminal truncat
39	105	7.9	2332	22	AAB50465	Human factor VIII
40	105	7.9	2332	23	AAD79869	Human factor VIII
41	105	7.9	2332	23	AAU79870	Human factor VIII
42	105	7.9	2332	23	AAU79871	Human factor VIII
43	105	7.9	2332	23	AAU79872	Human factor VIII
44	105	7.9	2332	23	AAU79873	Human factor VIII
45	105	7.9	2332	23	AAU79874	Human factor VIII

ALIGNMENTS

RESULT	1
AAy92240	
ID	AAy92240 standard; Protein; 254 AA.
XX	
AC	AAy92240;
XX	
DT	10-AUG-2000 (first entry)
XX	
DE	Human bone marrow-derived serum protein 2.
XX	
KW	Bone marrow-derived serum protein; BMDSP-2; MSE55; cytostatic; anti-HIV;
KW	antithrombotic; antihypertensive; antidiabetic; anti-inflammatory; antitumor;
KW	antiatherosclerotic; anti-infective; antiviral; antiparasitic.
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers 1..218 /note= "homologous with MSE55 N-terminal region"
FT	Region
FT	26..68 /label= glycine-rich
FT	/note= "homologous to MSE55"
FT	Modified-site
FT	88 /note= "potential N-glycosylation site"
FT	Modified-site
FT	89 /note= "potential casein kinase II phosphorylation site"
FT	Modified-site
FT	130 /note= "potential casein kinase II phosphorylation site"
FT	Modified-site
FT	158 /note= "potential casein kinase II phosphorylation site"
FT	Modified-site
FT	162 /note= "potential N-glycosylation site"
FT	Modified-site
FT	191

FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT 209
FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT 218
FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT 223
FT Binding-site /note= "potential casein kinase II phosphorylation site"
FT 235..245
FT /label= metal_binding_motif
FT /note= "homologous to MSE55"
FT 241
FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT
PN WO200020588-A2.
PD 13-APR-2000.
XX
XX
PF 01-OCT-1999; 99WO-US22908.
XX
XX
PR 02-OCT-1998; 98US-0165621.
XX
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Corley NC, Guegler KJ, Lu DAM;
XX
DR WPI; 2000-303775/26.
DR N-PSDB; AAA09155.
XX
XX
PT Purified polypeptide for treating or preventing disorders associated
PT with decreased expression or activity of bone marrow-derived serum
PT proteins
XX
XX
PS Claim 1; Page 69; 72pp; English.
XX
XX
CC Human bone marrow-derived serum proteins (BMDSP) 2 has chemical and
CC structural similarity with MSE55. BMDSP-1 and BMDSP-2 are useful for
CC treating or preventing a disorder associated with decreased expression
CC or activity of BMDSP.
CC Antagonists of BMDSP are useful for treating or preventing a disorder
CC associated with increased expression or activity of bone marrow-derived
CC serum proteins. The disorders include cancers (melanoma, adenocarcinoma,
CC sarcoma), immune disorders (acquired immunodeficiency syndrome (AIDS),
CC asthma, atherosclerosis, Crohn's disease, bronchitis, multiple sclerosis,
CC osteo- and rheumatoid arthritis), viral infections, parasitic infections
CC (schistosoma, tapeworm), and vascular disorders (arteriosclerosis,
CC hypertension, vasculitis).
XX
XX
SQ Sequence 254 AA;
Query Match 100.0%; Score 1326; DB 21; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.5e-126;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC/PL

RESULT 2
ABG25106
ID ABG25106 standard; Protein; 538 AA.
XX
XX
AC ABG25106;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #25097.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS89293.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 20; SEQ ID NO 55465; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 538 AA;
Query Match 42.4%; Score 562.5; DB 22; Length 538;
Best Local Similarity 66.1%; Pred. No. 3e-48;
Matches 115; Conservative 10; Mismatches 18; Indels 31; Gaps 2;

RESULT 5
AAM53373
ID AAM53373 standard; Protein; 137 AA.
XX
AC AAM53373;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 25478.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 25478; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 137 AA;
XX
Query Match 19.3%; Score 256.5; DB 22; Length 137;
Best Local Similarity 53.8%; Pred. No. 5.6e-18;
Matches 56; Conservative 10; Mismatches 35; Indels 3; Gaps 2;
OY 1 MPAKTPYLYKAAANNKGGKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 35 MSTKVPYLYK-RGSRKGKKEKLRDILSSDMISPLGDFRHTIHIGSGGSDMFGDISFLQ 93
OY 61 GNYELLPGNQEKA--HLGQFPGHNEFFRANSTSDSVFTETPSV 102
DB 94 GKFHLLPGTMEGPEEDGTFLDLPQFTRTATVCGRELDPGSPSL 137
RESULT 6
AAM13611
ID AAM13611 standard; Protein; 137 AA.
XX
AC AAM13611;
XX
DT 12-OCT-2001 (first entry)
XX

DE Peptide #45 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 18437; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 137 AA;
XX
Query Match 19.3%; Score 256.5; DB 22; Length 137;
Best Local Similarity 53.8%; Pred. No. 5.6e-18;
Matches 56; Conservative 10; Mismatches 35; Indels 3; Gaps 2;
OY 1 MPAKTPYLYKAAANNKGGKFKLRDILSPDMISPLGDFRHTIHIGKEGQHDVFGDISFLQ 60
DB 35 MSTKVPYLYK-RGSRKGKKEKLRDILSSDMISPLGDFRHTIHIGSGGSDMFGDISFLQ 93
OY 61 GNYELLPGNQEKA--HLGQFPGHNEFFRANSTSDSVFTETPSV 102
DB 94 GKFHLLPGTMEGPEEDGTFLDLPQFTRTATVCGRELDPGSPSL 137
RESULT 7
AAM26008
ID AAM26008 standard; Protein; 137 AA.
XX
AC AAM26008;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #45 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX

[illegible]

```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 27; SEQ ID NO 10101; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 137 AA;
QY
Query Match 19.3%; Score 256.5; DB 22; Length 137;
Best Local Similarity 53.8%; Pred. No. 5.6e-18;
Matches 56; Conservative 10; Mismatches 35; Indels 3; Gaps 2;
QY 1 MPAKPTIYLKANNKKGKKFKLDILSPDMISPLGDFRHTHTHGEGQHDVFGDISFLQ 60
1 | | | | | : : | | | | | | | | | | | | | | | | : | | | | |
Db 35 MSIRKVPYIYK-RGSRKKGKKEKRLDLSSDMISPLGDFRHTHTHGSGGSDMFGDISFLQ 93
QY 61 GNYELLPGNQEKA--HLGQFPGHNEFFRANSTSDSVFTETPPSPV 102
1 : : | | | | : : | | : : | | : : | | : : | | :
Db 94 GKRHLLPGTMEVGEPEEDGTFDLPQFTRTATVCGRELPGDPSPL 137
RESULT 9
ABG35381
ID ABG35381 standard; Peptide; 137 AA.
XX
AC ABG35381:
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 25046.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.

```

PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 25046; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenar syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 137 AA;
 Query Match 19.3%; Score 256.5; DB 23; Length 137;
 Best Local Similarity 53.8%; Pred. No. 5.6e-18;
 Matches 56; Conservative 10; Mismatches 35; Indels 3; Gaps 2;
 QY 1 MPAKPTLYLKANNKKGKFKLRLDILSPDMISPLGDFRHTIHGKGGHDFVGDISFLQ 60
 DB 35 MSTKVPYLYLK-RGSRKKGKKEKRLDILSSDMISPLGDFRHTIHGSGGSDMGDISFLQ 93
 QY 61 GNYELLPGNQEKA--HLGQFPGHNEFFRANSTSDSVFTETPPSPV 102
 DB 94 GKFLHLPGTWVEGPEEDGTFDLFPQFTRTATVCGRELDPGSPSL 137

RESULT 10
 ID AAE01862
 AC AAE01862 standard; Protein; 349 AA.
 XX
 AC AAE01862;
 XX
 DT 31-JUL-2001 (first entry)
 DE Mouse zmsel protein.
 XX
 KW Mouse; zmsel protein; Cdc42/Rac interactive binding protein; CRIB;
 KW Wiskott-Aldrich Syndrome; cancer; tumour; invasion; metastasis; asthma;
 KW digestion; actin polymerisation; cytoskeletal reorganisation; arthritis;
 KW testicular function; muscle inflammation; inflammatory bowel disease;
 KW diverticulitis; male infertility; male contraceptive agent; myocarditis;
 KW spermatogenesis; sperm capacitation; reperfusion ischaemia; psoriasis;
 KW melanoma; atherosclerosis; pelvic inflammatory disease; PID; eczema;
 KW scleroderma; cytoskeletal; vasotropic; dermatological; gene therapy.
 XX
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Domain 1..145 /note= "Conserved N-terminal domain"
 FT Domain 27..41 /label= CRIB_motif
 FT Domain 145..329 /note= "Variable C-terminal domain"
 FT Region 329..350 /note= "Highly conserved C-terminal tail"
 XX
 PN WO200134803-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US30945.
 XX
 PR 10-NOV-1999; 99US-0438564.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Whitmore TE;
 XX
 DR WPI; 2001-335928/35.
 DR N-PSDB; AAD05673.
 XX
 PT Novel human CRIB protein, zmsel and polynucleotide encoding the
 PT protein, for detecting human chromosomal abnormalities and for treating
 PT cancer, cardiovascular and inflammatory conditions -
 XX
 PS Claim 21; Fig 2; 132pp; English.
 XX
 CC The present invention relates to DNA and protein for zmsel, a novel
 CC human Cdc42/Rac interactive binding (CRIB) protein. CRIB proteins are
 CC implicated in human disease such as Wiskott-Aldrich Syndrome. Zmsel
 CC modulators are useful for modulating tumour cell motility, invasion and
 CC metastasis, gene transcription, contractility of various tissues, actin
 CC polymerisation and cytoskeletal reorganisation, digestion, testicular
 CC function and fertility. Zmsel sequence and its modulators are useful for
 CC treating cancer, inflammatory heart or cardiovascular conditions, muscle
 CC inflammation, inflammation during and after surgery, arthritis, asthma,
 CC inflammatory bowel diseases or diverticulitis, myocarditis, scleroderma,
 CC atherosclerosis, pelvic inflammatory disease (PID), eczema and other
 CC inflammatory diseases, male infertility or as male contraceptive agents
 CC and for modulating spermatogenesis and sperm capacitation. zmsel and
 CC anti-zmsel antibodies are useful in diagnosing inflammatory diseases,
 CC such as reperfusion ischaemia, psoriasis, arthritis, melanoma and other
 CC inflammatory diseases, male reproductive cancers such as prostate and
 CC testicular cancers. Zmsel polynucleotide sequences are useful as probes
 CC or primers for detecting human chromosomal abnormalities. zmsel sequence
 CC is used in gene therapy. The present amino acid sequence is mouse
 CC zmsel protein.

FT Modified-site /note="Hydrophilic region"
FT 295..297
FT /note="Phosphorylation site"
FT 303..305
FT /note="Phosphorylation site"
FT 337..356
FT /note="Highly conserved C-terminal tail"
FT 346..351
FT /note="Hydrophilic region"
FT 347..352
FT /note="Hydrophilic region"
FT 348..353
FT /note="Hydrophilic region"

WO200134803-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US30945.

10-NOV-1999; 99US-0438564.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Whitmore TE;

WPI; 2001-335928/35.

N-PSDB; AAD05671.

Novel human CRIB protein, zmsel and polynucleotide encoding the protein, for detecting human chromosomal abnormalities and for treating cancer, cardiovascular and inflammatory conditions

Claim 11; Fig 2; 132pp; English.

The present invention relates to DNA and protein for zmsel, a novel human Cdc42/Rac interactive binding (CRIB) protein. CRIB proteins are implicated in human disease such as Wiskott-Aldrich Syndrome. Zmsel modulators are useful for modulating tumour cell motility, invasion and metastasis, gene transcription, contractility of various tissues, actin polymerisation and cytoskeletal reorganisation, digestion, testicular function and fertility. Zmsel sequence and its modulators are useful for treating cancer, inflammation heart or cardiovascular conditions, muscle inflammation, inflammation during and after surgery, arthritis, asthma, inflammation, bowel diseases or diverticulitis, myocarditis, scleroderma, atherosclerosis, pelvic inflammatory disease (PID), eczema and other inflammatory diseases, male infertility or as male contraceptive agents and for modulating spermatogenesis and sperm capacitation. Zmsel and anti-zmsel antibodies are useful in diagnosing inflammatory diseases, such as reperfusion ischaemia, psoriasis, arthritis, melanoma and other inflammatory diseases, male reproductive cancers such as prostate and testicular cancers. Zmsel polynucleotide sequences are useful as probes or primers for detecting human chromosomal abnormalities. zmsel sequence is used in gene therapy. The present amino acid sequence is human zmsel protein. The zmsel gene is located at the 17q24.1 region of chromosome 17.

Sequence 356 AA;

Query Match 16.8%; Score 222.5; DB 22; Length 356;
Best Local Similarity 28.3%; Pred. No. 7e-14;
Matches 77; Conservative 37; Mismatches 91; Indels 67; Gaps 11;

QY 6 PIYKAANNKKGKFKLRDILSPDMISPLGDFRHTIHGEGHDFGDISFLGNYEL 65
DB 2 PILKQVSSSVHSKRSRADLTAEMLISAPLGDFRHTMHVGRAG--DAFGDTSLNSK--- 56
QY 66 LRGNOEKAHLGQFPCHNEF-----FRANSTSDSV-----FTETPSVYKLN 105
DB 57 -AGEPDGESLDEQSSSSSKRSLSRKFRGSKRSQSVTRGERQRDMGLRDSALFVKN 115
QY 106 AISLPTIGSGQALMLPLSPVTENSKQESFGPAKLPR-LSCPEVMEKAQEKSLLENGT 164

Db 116 AMSLPQUNEKA-----AEKGTSLPKLSLSSSPV--KKANDGEG----- 152
QY 165 VHGDTSWSSGSSASQSGRDSHSS--LSEQYPDWPAEDMFDPCELIKGTKSEE 222
Db 153 ---GDEAGTEEAVERRNGAAGPHSPDPLDEQ---AFGDLTDLPPVPKATYGLKHAE- 204
QY 223 SLSDLTGSLSLQDLGPPSLDEVLMVMDKNK 254
Db 205 -----SIMSFHIDGPPSMLGDLSIMDKEE 229

RESULT 13

AAB95088

AAB95088 standard; Protein; 356 AA.

AAB95088;

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:17019.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Claim 8; SEQ ID 17019; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1356 AA;

Query Match 9.8%; Score 130.5; DB 22; Length 1356;
Best Local Similarity 24.3%; Pred. No. 0.0013;
Matches 63; Conservative 34; Mismatches 103; Indels 59; Gaps 13;

QY 1 MPAKTPYLYKAANNKKGKFKLRDILSPDMISPLGDFRHTIHGKEGQHDVFGDISFL- 59
DB 462 LPSSTRDSFSRVSDFHRISKRLR---TEMISKPQNDKHTGHVIGIDGA--TFGDIAFLG 515
QY 60 -QGNVELLPNGNEKAHLGQFPGHNEFFRANSTSDSVFTETP---SPVLKNAISLPTIGS 115
DB 516 SSQNYNHVP-----KQIVTPYKPSEDI-EQTPLLLPPTPTSPDSLQTAGS- 559
QY 116 QALMLPLSPVTFNS---KQESFGPAKLPRLSCEPVMEKAQEKSSLENGTV----HQ 167
DB 560 -----YFPEGANSGAMGTSMNPTFIPSAEHTPKLIATNGOSSSDFASGSTNPFDPNR 612
QY 168 GDTSWGSSGSASQSSQGRDSSSSLSSEQYPDWPADMFDPHTPCCELLIKGKTKEESLSDL 227
DB 613 GDDEL-EFGGLHNYGADGKSVHSET-----GW-----RPTSRISIVDDPHHEYHEISDD- 657
QY 228 TGSLLSLQILDGPSLLDEV 246
DB 658 --EIAADKLDGFPSSLDEI 674

Search completed: April 23, 2003, 11:01:08
Job time : 56.6516 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 10:59:20 ; Search time 13.5328 Seconds
(Without alignments)
552.246 Million cell updates/sec

Title: US-09-806-276A-2

Perfect score: 1326
Sequence: 1 MPAKPTPIYKAAANNKKKKF.....QLDLGPSLDEVLYNVDKNK 254

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	7.9	2332	1	US-07-864-004B-4 Sequence 4, Appl1
2	105	7.9	2332	1	US-08-251-937A-4 Sequence 4, Appl1
3	105	7.9	2332	1	US-08-212-133A-2 Sequence 2, Appl1
4	105	7.9	2332	1	US-08-276-594A-2 Sequence 2, Appl1
5	105	7.9	2332	1	US-08-474-503-2 Sequence 2, Appl1
6	105	7.9	2332	2	US-08-670-707A-2 Sequence 2, Appl1
7	105	7.9	2332	4	US-09-037-601-2 Sequence 2, Appl1
8	105	7.9	2332	4	US-09-324-867-3 Sequence 3, Appl1
9	105	7.9	2332	4	US-09-315-179-2 Sequence 2, Appl1
10	105	7.9	2332	4	US-09-523-656-2 Sequence 2, Appl1
11	105	7.9	2332	5	PCT-US93-03275-4 Sequence 4, Appl1
12	105	7.9	2332	5	PCT-US94-13200-2 Sequence 2, Appl1
13	105	7.9	2351	1	US-08-121-202-2 Sequence 2, Appl1
14	105	7.9	2351	1	US-08-366-851A-2 Sequence 2, Appl1
15	105	7.9	2351	6	5171844-2 Patent No. 5171844
16	105	7.9	2351	6	5422260-1 Patent No. 5422260
17	94	7.1	416	2	US-09-211-930-11 Sequence 11, Appl1
18	94	7.1	416	3	US-09-340-993-11 Sequence 11, Appl1
19	94	7.1	416	4	US-09-468-442-11 Sequence 11, Appl1
20	88	6.6	933	3	US-08-293-728-2 Sequence 2, Appl1
21	88	6.6	933	4	US-09-421-868-2 Sequence 2, Appl1
22	86	6.5	716	4	US-09-219-983A-20 Sequence 20, Appl1
23	86	6.5	1079	3	US-09-058-489-22 Sequence 22, Appl1
24	86	6.5	1142	2	US-08-993-118-7 Sequence 7, Appl1
25	86	6.5	1142	3	US-08-845-528C-7 Sequence 7, Appl1
26	86	6.5	1142	4	US-09-061-709-2 Sequence 2, Appl1
27	86	6.5	1240	3	US-09-058-489-23 Sequence 23, Appl1

28	86	6.5	1347	3	US-09-058-489-24	Sequence 24, Appl1
29	85.5	6.4	966	1	US-08-571-758-2	Sequence 2, Appl1
30	85.5	6.4	966	1	US-08-909-984A-2	Sequence 2, Appl1
31	85.5	6.4	966	1	US-08-909-983-2	Sequence 2, Appl1
32	85	6.4	649	2	US-08-149-097D-37	Sequence 37, Appl1
33	83.5	6.3	224	4	US-08-944-483-34	Sequence 34, Appl1
34	83.5	6.3	462	3	US-08-875-944B-5	Sequence 5, Appl1
35	83.5	6.3	462	4	US-09-116-049-4	Sequence 21, Appl1
36	83.5	6.3	1037	4	US-09-428-711A-21	Sequence 11, Appl1
37	83	6.3	478	1	US-08-127-499A-11	Sequence 11, Appl1
38	83	6.3	478	1	US-08-482-847-11	Sequence 11, Appl1
39	83	6.3	512	3	US-08-463-210-8	Sequence 8, Appl1
40	83	6.3	512	4	US-09-124-900-2	Sequence 8, Appl1
41	83	6.3	2137	4	US-09-134-001C-4463	Sequence 4463, Ap
42	82.5	6.2	1637	4	US-09-718-692-2	Sequence 2, Appl1
43	82.5	6.2	1637	4	US-09-718-852-2	Sequence 2, Appl1
44	82.5	6.2	1637	4	US-09-718-815-2	Sequence 2, Appl1
45	82	6.2	500	1	US-08-375-510-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4
Query Match 7.9%; Score 105; DB 1; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

Query Match 7.9%; Score 105; DB 1; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPPIYKANNKKKKFKRLDILSPDMIS-----PPLGDFRHTIHIGKEGQHDVFGDISF- 58
DB 799 TPHGLSLDLQEAKEYETFSDDPSGAIDSNNSLSEMTFRPQLH-----HS--GDMVFT 850

QY 59 -----LQNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFTETPSVYL----KNAIS 108
DB 851 PESGLRLNEKL--GTTATELKKLD-----FKVSSSTNNLISTIPSDNLAAGTDNTSS 903

QY 109 LPTIGGSQALMLPLSPVTENSKQES--FGPAKLPRL--SCEPVMEKAEKQESSLLENGTV 165
DB 904 LG-----PPSMPVHYDSQDLDTTLFGKKSSPLTESGGLSLSEENNSKLLSGLM 953

QY 166 HQGDTWSSSGSASQSS--QGRDSHSSSLSEQYPPDPAEDMFDHPTPCELIKGKTKSEE 222
DB 954 NSQESSWGKNVSTESGRLFKGKRAHGPAALL-----TKDNALFKVISISLKTNTKTSNN 1006

QY 223 SLSD---LTGSLSLQDLGPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLIENSPSVQNLIESDTEFKKVTPLIHDMMLMDKN 1055

RESULT 4
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-276-594A-2
Query Match 7.9%; Score 105; DB 1; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPPIYKANNKKKKFKRLDILSPDMIS-----PPLGDFRHTIHIGKEGQHDVFGDISF- 58
DB 799 TPHGLSLDLQEAKEYETFSDDPSGAIDSNNSLSEMTFRPQLH-----HS--GDMVFT 850

QY 59 -----LQNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVFTETPSVYL----KNAIS 108
DB 851 PESGLRLNEKL--GTTATELKKLD-----FKVSSSTNNLISTIPSDNLAAGTDNTSS 903

QY 109 LPTIGGSQALMLPLSPVTENSKQES--FGPAKLPRL--SCEPVMEKAEKQESSLLENGTV 165
DB 904 LG-----PPSMPVHYDSQDLDTTLFGKKSSPLTESGGLSLSEENNSKLLSGLM 953

QY 166 HQGDTWSSSGSASQSS--QGRDSHSSSLSEQYPPDPAEDMFDHPTPCELIKGKTKSEE 222
DB 954 NSQESSWGKNVSTESGRLFKGKRAHGPAALL-----TKDNALFKVISISLKTNTKTSNN 1006

QY 223 SLSD---LTGSLSLQDLGPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLIENSPSVQNLIESDTEFKKVTPLIHDMMLMDKN 1055

RESULT 5
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-08-474-503-2

QY 5 TPITYLKANNKKKKFKLRDILSPDMIS-----PLGDFRHTIHIGKEGHVFGDISF- 58
|||:::| ||| : ||:|
Db 799 TPHGLSLDLQEAKEYETFSDDPSGAIDSNNSLSEMTHERPOLH-----HS--GDMVET 850

OY	109	LPTIGSQALMLPLSPVTENKQES--FGPAKLPR-LSCPEVMEKAQEKSSILENGTV	165
		: : : :	
Dd	904	LG-----PPSMFVHYDSQLDLTLLFGKKSSPLTSGGPLSLSEFNNDSKLLESGIM	953

```

1007 SATNRKTHIDGP--SLIENSPSVQNILLESDTEFFKVTPLIHDRMLMDKN 1055

```

Sequence 2, Application US/08670707A
Patent No. 5859204

```

; GENERAL INFORMATION:
;
; APPLICANT: Lollar, John S.
;
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;
; NUMBER OF SEQUENCES: 40
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESS:

```

ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303

```

;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996

CLASSIFICATION: 435
PRIOR APPLICATION DATA

APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994

ENJOY APPLICATION DATA:
 APPLICATION NUMBER: US 08/212,133
 FILING DATE: 1994

FILING DATE: 11-MAR-1994
 PRIOR APPLICATION DATA:
 ADDITIONAL NUMBER: 11-11-11

APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:

NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 37 804

IDENTIFICATION NUMBER: 21,894
 REFERENCE/DOCKET NUMBER: 75-95F
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS.

LENGTH: 2332 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: not relevant

MOLECULE TYPE: protein

Query Match	7.98;	Score 105;	DB 2;	Length 2332;
Best Local Similarity	23.08;	Pred. No. 0.12;		
Matches 67;	Conservative 46;	Mismatches 102;	Indels 76;	Gaps 15;

851 PESGLQRLNEKL-GTTAATELKKLD-----FKVSYSSNNLLISTIPSDNLAAGTDNTSS 903

[illegible]

Dbb

954 NSQESSWCKNVSTESGRLEFGKRAHGPAIL-----TKRDALKEVSTSIIRPMYDGVY 100

QY	223	SLSD----	LTGSLSLQDIDGPSLLDEVN-----	WARDN 252	
					1000

```

Ddb      1007 SATNRKTHIDGP--SLIENSFSVWQNLLESDTEFKVYTPLIHDMMDKN 105
          |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
RESULT 7  001 0 008-00-007 001

```

```

; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII

```

NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado

COUNTRY: USA
ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentInRelease #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 70 202 4000 1

AFFIDAVIT NUMBER: WO PCIT/US94/132000
 FILING DATE: 15-NOV-1994
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994

RECEIVED DATE: 11 MAR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: IIS 07/854 004

FILE NUMBER: 05-011884, 004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:

NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33 878

REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/499-8080

```

; INFORMATION FOR SEQ ID NO: 2 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

Query Match          7.9%; Score 105; DB 4; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

   5 TPVILKANNKKGGKKFKLRDILSPDMIS-----PPLGDFRHTIHIGKEGHDFGDISEF- 58
       ||| |:::|| | || | | | | | | | | | | | | | | | | | | | | | |
Db    799 TPHGLSLSDLOEAKYETFSDDPSGAIDSNNSLSMEWTHFRQQLH-----HS--GDMVF 850

   59 -----LQGNYELLPGNQEKAHLGQFPGHNEFFRANSTDSVTETPPVL----KNALS 108
           |:| |::|| | | | | | | | | | | | | | | | | | | | | | |
Db     851 PEGSGLRLNLEKL-GITAAATELKLD-----FKVSSTNNLISTIPSDNLAGTDNTSS 903

   109 LPTIGGSQAALMLPLSPVTFNSKOES-FGPAPLPRL-SCEPVMEKAQOEKSSLLENCTV 165
         | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     904 LG-----PPMPVHYDSQDLDTTLFGKKSSPLTESGGPLSLEENNDSKLESGLM 953

   166 HQGDTSWGSSGSASQS---QGRDSSHSSLSSEQYPDWPAEDMEDHPTECELIKGTSEE 222
         ::||| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db     954 NSQEWSWGKNVSTSGRLFKGKRAHGALL-----TKDNALFKVISISLTKTNTSNN 1006

   223 SLSD---LTGSLSLQDLDGPSLLDEVLN-----VMCKN 253
        |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db    1007 SATNRKTHIDGP-SLIIEGPSVWQNILESDTEFKKYTPPLIHDRIMMDKN 1055


RESULT 8
US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
; APPLICANT: Lillycrap, David.
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632ley, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.0010002/JAG/BJD
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-059
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-324-867-3

Query Match          7.9%; Score 105; DB 4; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

   5 TPVILKANNKKGGKKFKLRDILSPDMIS-----PPLGDFRHTIHIGKEGHDFGDISEF- 58
       ||| |:::|| | || | | | | | | | | | | | | | | | | | | | | | |
Db    799 TPHGLSLSDLOEAKYETFSDDPSGAIDSNNSLSMEWTHFRPQLH-----HS--GDMVFT 850
```

```
QY      59 -----LOGNYELLPGNOEKAHLGOFPGHNEFFRANSTSDSVTETPPVVL-----KNAIS 108  
        | : | : | |         | : : ||| :: : | | |       |  
Db     851 PESGLRLRNKL--GTTAATELKKLD----FKVYSTSNMLISTIPSDNLAAGTDNTSS 903  
  
QY     109 LPTIGGSQALMPLPSEVTENSKOES--FGPAKLPRL-SCEPVMEKEAQEKSSLLENQTV 165  
        |           | | | ::| : : | | | : | : | | | | : |  
Db     904 LG-----PPSMFVHYDSQLDITTLFGKKSPLTSGGPLSLSEENDSKILLESGLM 953  
  
QY     166 HQGDTSWSSGSASQS---QGRDSSHSSLSSEQYPMDAEDMFDPTELCelikgktksee 222  
        | : ||| : | : : | | : : | : | : | : | : | : |  
Db     954 NSQESSWGKNVSSTESGRLEKGRAHPALL-----TKDNALFKVISLSLKTNKTSNN 1006  
  
QY     223 SLSD---LTGSLSLQLDLGPSSLDEVLN-----VMCKN 253  
        | : : | : | | : : | : | : | : | : | : | : | : |  
Db    1007 SATNRKTHIDGP--SLLIENSFWQNILLESDTEEFKVYLP LIHDRMLMDKN 1055
```

```

RESULT 9
US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
;
; GENERAL INFORMATION:
;
; APPLICANT: Lollar, John S
;
; TITLE OF INVENTION: Modified Factor VIII
;
; FILE REFERENCE: 75-95H
;
; CURRENT APPLICATION NUMBER: US/09/315,179
;
; CURRENT FILING DATE: 1999-05-20
;
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
;
; EARLIER FILING DATE: 1998-03-10
;
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
;
; EARLIER FILING DATE: 1996-06-26
;
; EARLIER APPLICATION NUMBER: PCT/US97/11155
;
; EARLIER FILING DATE: 1997-06-26
;
; EARLIER APPLICATION NUMBER: PCT/US94/13200
;
; EARLIER FILING DATE: 1994-11-15
;
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
;
; EARLIER FILING DATE: 1994-03-11
;
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
;
; EARLIER FILING DATE: 1992-04-07
;
; NUMBER OF SEQ ID NOS: 40
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 2
;
; LENGTH: 2332
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-09-315-179-2

```

Query Match	7.98;	Score 105;	DB 4;	Length 2332;
Best Local Similarity	23.08;	Pred. No. 0.12;		
Matches 67;	Conservative 46;	Mismatches 102;	Indels 76;	Gaps

QY	5	TPVYLKANNKKKKFKLRLLISPDMTS-----PPLGDFRHTIHIGKEGHDFEGDISF-	58
		: : : : : :	
Db	799	TPHGLSLSDLQEAKEYETFSDDPSGAIDSNNSLSEMTHERPQLH-----HS--GDWVET	850
		: : : : : :	
QY	59	-----LQGNVELLPGNQEKALIGQEPGHNEFFRANSTSDSVETETPSVL---KNAIS	108
		: : : : : : : : :	
Db	851	PESGLQLRLNEKL--GTAAATELKKLD-----FKVSSFTSNMLISTIPSDNLAAGTDMISS	903
		: : : : : : : : :	
QY	109	LPTIGGSQALMLPLSPVTFNSKQES--FGPAKLPRL--SCPEVMEKAQEKSSLLENGTY	165
		: : : : : : : : : : :	
Db	904	LG-----PPSMFVPHIDSQLDITTLFGKKSSPLTESGCPLSLSEENNDSKLLESLGM	953
		: : : : : : : : : : : :	
QY	166	HOGDTSWSSSGSASOSS---OGRDSSHSSLSLSEQYPDWPAEDMFDPPTPCFLIKGKTSEE	222
		: : : : : : : : : : : : :	
Db	954	NSQESSWGKNVSSSTESGRLEFGKRAHGAPALL-----TRDNALLFKVISISLTKTNTISNN	1006
		: : : : : : : : : : :	
QY	223	SLSD---LTGSLSLQLDLGPSLLDEVLN-----VMDKN	253
		: : : : : : : : : : :	
Db	1007	SATNRKTHIDGP--SLIENSPSVQONILESDTEFFKKVTPPLIHDRMLMDKN	1055
		: : : : : : : : : : :	

RESULT 10
US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-951
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

Query Match 7.9%; Score 105; DB 4; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;
QY 5 TPVYKANNKGGKFKLRDILSPDMIS-----PPLGDFRHTIHIGKEQHDFGDISF- 58
DB 799 TPHGLSLDLQAKYETFPDPSGAIDSNNSISEMTHFRPQLH-----HS--GDMVFT 850
QY 59 -----LQGNVELLPNGQEAHLGQFPGHNEFFRANSTSDSVFTETPSPVL---KNAIS 108
DB 851 PESGLQLRLNEKI--GTTATTELKLD-----FKVSTSNLITIPSDNLAAGTDNTSS 903
QY 109 LPTIGGSQALMLPLSPVTFNSKQES--FGPAKLPR--SCEPVMEKAEKSSILENGTV 165
DB 904 LG-----PPSMVHYDSQDLDTLFGKSSPLTESGGLSEENNDKSKLLESGLM 953
QY 166 HOGDTWSSGSSASQSS---QGRDSSSSLSSEQYPDWPAEDMFDPHPTPCELIKGKTSEE 222
DB 954 NSQSSWGKNVSTESGRLEFKGKRAHGPALL-----TKDNALFKVSIISLKTNTKTSNN 1006
QY 223 SLSD---LTGSLSLQDLGLPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLIENSPSVQWNLIESDTEFKKVTPLIHDMMDKN 1055

RESULT 11
PCT-US93-03275-4
; Sequence 4, Application PC/TUS9303275
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03275
; FILING DATE: 19930407
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864004

; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 106PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
PCT-US93-03275-4

Query Match 7.9%; Score 105; DB 5; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;
QY 5 TPVYKANNKGGKFKLRDILSPDMIS-----PPLGDFRHTIHIGKEQHDFGDISF- 58
DB 799 TPHGLSLDLQAKYETFPDPSGAIDSNNSISEMTHFRPQLH-----HS--GDMVFT 850
QY 59 -----LQGNVELLPNGQEAHLGQFPGHNEFFRANSTSDSVFTETPSPVL---KNAIS 108
DB 851 PESGLQLRLNEKI--GTTATTELKLD-----FKVSTSNLITIPSDNLAAGTDNTSS 903
QY 109 LPTIGGSQALMLPLSPVTFNSKQES--FGPAKLPR--SCEPVMEKAEKSSILENGTV 165
DB 904 LG-----PPSMVHYDSQDLDTLFGKSSPLTESGGLSEENNDKSKLLESGLM 953
QY 166 HOGDTWSSGSSASQSS---QGRDSSSSLSSEQYPDWPAEDMFDPHPTPCELIKGKTSEE 222
DB 954 NSQSSWGKNVSTESGRLEFKGKRAHGPALL-----TKDNALFKVSIISLKTNTKTSNN 1006
QY 223 SLSD---LTGSLSLQDLGLPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLIENSPSVQWNLIESDTEFKKVTPLIHDMMDKN 1055

RESULT 12
PCT-US94-13200-2
; Sequence 2, Application PC/TUS9413200
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13200
; FILING DATE: 15-NOV-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106CIP(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
PCT-US94-13200-2

Query Match 7.9%; Score 105; DB 5; Length 2332;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPYLLKANNKKKKRRLDILSPDMIS-----PPLGDFRHTIHIGEGQHDVFGDISF- 58
DB 799 TPHGSLSDLOEAKYETFDSPGAIDSNNSLSEMTFRPQLH-----HS--GDMVFT 850
QY 59 -----LQGNVELLPNGQEKALGQFPNGNEFFRANSTSDSVFTETPSVYL---KNAIS 108
DB 851 PESGLQLRLNEKL--GTTAATELKKLD-----FKVSSSTNNLISTIPSDNLAAGTDNTSS 903
QY 109 LPTIGSQALMLPLSPVTFNSKQES--FGPAKLPR-L-SCEPVMEKAKQEKSSILENGTV 165
DB 904 LG-----PPSMFVHYDSQDLDTLFGKSSPLTESGGLSLSENNDSKLLSGLM 953
QY 166 HQGDTWSGSSGSASQSS---QGRDSSSSLSSEQYPPDPAEDMFDHPTPCELIKGKTSEE 222
DB 954 NSQESSWGKNVSTESGRLFKGKRAHGAPALL-----TKDNALFKVSIISLKTNTKTSNN 1006
QY 223 SLSD----LTGSLSLQLDLGPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLIENSPSVWQNILESDETEFKVTPLIHDRMLMDKN 1055

RESULT 13

US-08-121-202-2
Sequence 2, Application US/08121202
Patent No. 5563045
GENERAL INFORMATION:
APPLICANT: Pittman, Debra
APPLICANT: Rehmentulla, Alnawaz
APPLICANT: Wozney, John M.
APPLICANT: Kaufman, Randal J.
TITLE OF INVENTION: CHIMERIC PROCOAGULANT PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,202
FILING DATE: 14-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Meinerdt, M. C.
REGISTRATION NUMBER: 31,544
REFERENCE/DOCKET NUMBER: GI 5195A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1210 X8574
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2351 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-121-202-2

Query Match 7.9%; Score 105; DB 1; Length 2351;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPYLLKANNKKKKRRLDILSPDMIS-----PPLGDFRHTIHIGEGQHDVFGDISF- 58
DB 818 TPHGSLSDLOEAKYETFDSPGAIDSNNSLSEMTFRPQLH-----HS--GDMVFT 869
QY 59 -----LQGNVELLPNGQEKALGQFPNGNEFFRANSTSDSVFTETPSVYL---KNAIS 108
DB 870 PESGLQLRLNEKL--GTTAATELKKLD-----FKVSSSTNNLISTIPSDNLAAGTDNTSS 922
QY 109 LPTIGSQALMLPLSPVTFNSKQES--FGPAKLPR-L-SCEPVMEKAKQEKSSILENGTV 165
DB 923 LG-----PPSMFVHYDSQDLDTLFGKSSPLTESGGLSLSENNDSKLLSGLM 972
QY 166 HQGDTWSGSSGSASQSS---QGRDSSSSLSSEQYPPDPAEDMFDHPTPCELIKGKTSEE 222
DB 973 NSQESSWGKNVSTESGRLFKGKRAHGAPALL-----TKDNALFKVSIISLKTNTKTSNN 1025
QY 223 SLSD----LTGSLSLQLDLGPSLDEVLN-----VMDKN 253
DB 1026 SATNRKTHIDGP--SLIENSPSVWQNILESDETEFKVTPLIHDRMLMDKN 1074

RESULT 14

US-08-366-851A-2
Sequence 2, Application US/08366851A
Patent No. 5681746
GENERAL INFORMATION:
APPLICANT: Bodner, Mordechai
APPLICANT: De Polo, Nicolas J.
APPLICANT: Hsu, David Chi-Tang
APPLICANT: Chang, Steven
TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Viagene, Inc.
STREET: 11055 Roselle Street
CITY: San Diego
STATE: California
COUNTRY: U.S.A.
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,851A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Chambers, Daniel M.
REGISTRATION NUMBER: 34,561
REFERENCE/DOCKET NUMBER: 930049.438
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 452-1288
TELEFAX: (619) 452-2616

Search completed: April 23, 2003, 11:04:02
Job time : 23.5328 secs

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-366-851A-2

Query Match 7.9%; Score 105; DB 1; Length 2351;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPVYLAANNKKKKFKLDILSPDMIS-----PPLGDFRHTIHIGKEGQHDVFGDISF- 58
DB 818 TPVGLSLSDLOEAKYETFSDDPSGAIDSNLSSEMTFRPQLH-----HS--GDMVFT 869
QY 59 -----LQGNVELLPNGQEKALGQFPGHNEFFRANSTSDSVFTETPSVYL---KNAIS 108
DB 870 PESGLQLRLNEKL--GTTAATELKKLD---FKVSTSNLSTIPSDNLAAGTDNTSS 922
QY 109 LPTIGSQALMLPLSPVTENSKQES--FGPAKLPRL--SCEPVMEKAEKSSLENGTV 165
DB 923 LG-----PPSMFVHYDSQLDITLFGKSSPLTESGGPLSLSEENDSKLLESGLM 972
QY 166 HGGDTSMGSSGSASQSS---QGRDSSSSLSSEQYPDWPAEDMFDPCELIKGTKSEE 222
DB 973 NSQESSWGKNVSTESGRLEKGRAGHPALL-----TKDNALFKVSI SLKTNKTSNN 1025
QY 223 SLSD---LTGSLSLQDLGSPSLDEVLN-----VMDKN 253
DB 1026 SATNRKTHIDGP--SLTIENSPSVWQNI LESDTEFKKVTPLIHDM LMDKN 1074

RESULT 15
5171844-2
; Patent No. 5171844
; APPLICANT: VAN OYEN, ALBERT J.J.; PANNEKOEK, HANS; VERBEET,
; MARTINUS P.; VAN LEEN, ROBERT W.
; TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY
; PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS
; AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/205,226
; FILING DATE: 10-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 2351
5171844-2

Query Match 7.9%; Score 105; DB 6; Length 2351;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPVYLAANNKKKKFKLDILSPDMIS-----PPLGDFRHTIHIGKEGQHDVFGDISF- 58
DB 818 TPVGLSLSDLOEAKYETFSDDPSGAIDSNLSSEMTFRPQLH-----HS--GDMVFT 869
QY 59 -----LQGNVELLPNGQEKALGQFPGHNEFFRANSTSDSVFTETPSVYL---KNAIS 108
DB 870 PESGLQLRLNEKL--GTTAATELKKLD---FKVSTSNLSTIPSDNLAAGTDNTSS 922
QY 109 LPTIGSQALMLPLSPVTENSKQES--FGPAKLPRL--SCEPVMEKAEKSSLENGTV 165
DB 923 LG-----PPSMFVHYDSQLDITLFGKSSPLTESGGPLSLSEENDSKLLESGLM 972
QY 166 HGGDTSMGSSGSASQSS---QGRDSSSSLSSEQYPDWPAEDMFDPCELIKGTKSEE 222
DB 973 NSQESSWGKNVSTESGRLEKGRAGHPALL-----TKDNALFKVSI SLKTNKTSNN 1025
QY 223 SLSD---LTGSLSLQDLGSPSLDEVLN-----VMDKN 253
DB 1026 SATNRKTHIDGP--SLTIENSPSVWQNI LESDTEFKKVTPLIHDM LMDKN 1074

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: April 23, 2003, 11:01:15 ; Search time 15.6148 Seconds
(without alignments)
1303.445 Million cell updates/sec

Title: US-09-806-276A-2
Perfect score: 1326
Sequence: 1 MPAKTPYIKANNKKKKF.....QLDLGPSLDELVDKKNK 254

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256.5	19.3	137	10	US-09-864-761-42650
2	109.5	8.3	938	9	US-10-029-217A-28
3	109.5	8.3	938	9	US-10-029-217A-29
4	105	7.9	2332	9	US-09-957-641-2
5	105	7.9	2332	9	US-10-187-319-2
6	105	7.9	2351	9	US-10-132-829-4
7	104	7.8	2344	10	US-09-815-242-12713
8	92	6.9	560	10	US-09-815-242-13057
9	88.5	6.7	1239	9	US-09-291-417-13
10	88	6.6	936	1	US-08-781-986A-5249
11	87	6.6	371	9	US-09-533-029-110
12	86	6.5	271	9	US-09-925-299-806
13	86	6.5	271	10	US-09-925-299-806
14	86	6.5	473	10	US-09-864-761-38321
15	86	6.5	525	10	US-09-814-550-2
16	86	6.5	617	10	US-09-817-676A-12
17	86	6.5	1031	10	US-09-815-242-10932
18	86	6.5	1142	9	US-10-085-108-7
19	86	6.5	1142	10	US-09-899-651-2

20	86	6.5	1175	10	US-09-771-161A-224	Sequence 224, App
21	86	6.5	1175	10	US-09-771-161A-225	Sequence 225, App
22	86	6.5	1175	10	US-09-771-161A-226	Sequence 226, App
23	86	6.5	1702	9	US-09-854-133-434	Sequence 434, App
24	86	6.5	1702	10	US-09-738-973-434	Sequence 434, App
25	85.5	6.4	741	9	US-09-738-626-4455	Sequence 4455, App
26	85.5	6.4	1237	9	US-10-108-605-211	Sequence 211, App
27	85.5	6.4	1394	9	US-10-108-605-213	Sequence 213, App
28	85.5	6.4	3256	10	US-09-919-172-98	Sequence 98, App
29	85	6.4	568	10	US-09-820-905-2	Sequence 2, App
30	85	6.4	589	10	US-09-820-905-4	Sequence 4, App
31	84	6.3	606	10	US-09-820-905-5	Sequence 5, App
32	84	6.3	1002	9	US-09-988-117-3	Sequence 3, App
33	84	6.3	1002	10	US-09-812-471-3	Sequence 3, App
34	84	6.3	1002	10	US-09-812-633-3	Sequence 3, App
35	84	6.3	1350	9	US-09-952-060-35	Sequence 35, App
36	83.5	6.3	247	10	US-09-923-779-154	Sequence 154, App
37	83.5	6.3	462	9	US-09-884-363-4	Sequence 4, App
38	83	6.3	363	9	US-10-003-035-37	Sequence 37, App
39	83	6.3	453	9	US-10-003-035-39	Sequence 39, App
40	83	6.3	599	9	US-10-003-035-59	Sequence 59, App
41	83	6.3	764	9	US-10-081-119-30	Sequence 30, App
42	82.5	6.2	227	10	US-09-864-761-36295	Sequence 36295, A
43	82.5	6.2	276	10	US-09-864-761-38306	Sequence 38306, A
44	82	6.2	1021	10	US-09-815-242-5471	Sequence 5471, App
45	82	6.2	1021	10	US-09-815-242-12544	Sequence 12544, A

ALIGNMENTS

RESULT 1
US-09-864-761-42650
Sequence 42650, Application US/098647761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-x-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42650
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC000353.21
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 39
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.91
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: SWISSPROT HIT: Q00587, EVALUATE 3.00e-07
OTHER INFORMATION: EST_HUMAN HIT: BF058866.1, EVALUATE 4.00e-76
OTHER INFORMATION: EST_HUMAN HIT: T75138.1, EVALUATE 3.00e-60
US-09-864-761-42650

Query Match 19.3%; Score 256.5; DB 10; Length 137;
Best Local Similarity 53.8%; Pred. No. 3.2e-16;
Matches 56; Conservative 10; Mismatches 35; Indels 3; Gaps 2;

QY 1 MPKPTIYLKANKKGGKFKLRDILSPDMISPLGDFRHTIHKEGQHDVFGDISFLQ 60
DB 35 MSTKVPITLK-RGSRKGGKKEKRLDLSMDISPLGDFRHTIHKSGGSDMGDISFLQ 93
QY 61 GNYELLPGNOEKA--HLGQFPGHNEFFRANSTSDSVFTETPSV 102
DB 94 GKFHLLPGTMEGPEEDGTFLPFGFTTATVCGRELPGSPSL 137

RESULT 2
US-10-029-217A-28
Sequence 28, Application US/10029217A
Patent No. US20020164735A1
GENERAL INFORMATION:
APPLICANT: OLSON, ERIC N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
TITLE OF INVENTION: NUCLEAR REGULATORY FACTOR
FILE REFERENCE: UTSD:695US
CURRENT APPLICATION NUMBER: US/10/029,217A
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 60/257,761
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 938
TYPE: PRT
ORGANISM: Homo sapiens
US-10-029-217A-28

Query Match 8.3%; Score 109.5; DB 9; Length 938;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 81; Conservative 31; Mismatches 106; Indels 117; Gaps 17;

QY 2 PAKPTIYLKANKKGGKFKLRDILSPDMISPLGDFRHTIHKEGQHDVFGDISFLQ 40
DB 227 PPTPIAVHAAVSKSLGSKNRHKKPKDKPKVKKLKHQYIPDQKAESPPMDSAY 286
QY 41 TIIHKEGQHDVFGDISFLQNGNYELLPGNOEK-----AHLGQFPGHNEFFRANSTS 91

DB 287 ARLLQOQO-----LFLQ--LQILSQOQOQOQHRFSTYLGMOAQLKEPNEQVARNPNS 336
QY 92 DSV-FTEPPSPVLKNAIS-----LPTIGSQO 117
DB 337 SSTPLSNTPLSPVKNFSGQGVSSFKPGPLPNDLKLKSELROQLRIRGLPVSGTKTA 396
QY 118 LMLPL-----LSPVTENSKQESFGPAKLPRLSCEPVEEKAQEKSSLE 161
DB 397 LMDRLRPFQDCSGNPVPNFGDITVT-----PYTPNTLPNY-----QSSSTALS 443
QY 162 NGTVHGDFTSGSSGASQSGRDSHSSSLSEQYPDWPAEDMFD--HPTCELIKGKTKS 220
DB 444 NGFYHFGSTS--SSPISPASSDL-SVAGSLPDTFND--ASPSFGLHPSVPVHVC-----T 493
QY 221 EESL-SDLTGLSLSLQLDLGPSSLDEVLNVMKKNK 254
DB 494 EESLMSLNGSVPSSELDGLDSEKDKML--VEKOK 526

RESULT 3
US-10-029-217A-29
Sequence 29, Application US/10029217A
Patent No. US20020164735A1
GENERAL INFORMATION:
APPLICANT: OLSON, ERIC N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
TITLE OF INVENTION: NUCLEAR REGULATORY FACTOR
FILE REFERENCE: UTSD:695US
CURRENT APPLICATION NUMBER: US/10/029,217A
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 60/257,761
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 938
TYPE: PRT
ORGANISM: Homo sapiens
US-10-029-217A-29

Query Match 8.3%; Score 109.5; DB 9; Length 938;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 81; Conservative 31; Mismatches 106; Indels 117; Gaps 17;

QY 2 PAKPTIYLKANKKGGKFKLRDILSPDMISPLGDFRHTIHKEGQHDVFGDISFLQ 40
DB 227 PPTPIAVHAAVSKSLGSKNRHKKPKDKPKVKKLKHQYIPDQKAESPPMDSAY 286
QY 41 TIIHKEGQHDVFGDISFLQNGNYELLPGNOEK-----AHLGQFPGHNEFFRANSTS 91
DB 287 ARLLQOQO-----LFLQ--LQILSQOQOQOQHRFSTYLGMOAQLKEPNEQVARNPNS 336
QY 92 DSV-FTEPPSPVLKNAIS-----LPTIGSQO 117
DB 337 SSTPLSNTPLSPVKNFSGQGVSSFKPGPLPNDLKLKSELROQLRIRGLPVSGTKTA 396
QY 118 LMLPL-----LSPVTENSKQESFGPAKLPRLSCEPVEEKAQEKSSLE 161
DB 397 LMDRLRPFQDCSGNPVPNFGDITVT-----PYTPNTLPNY-----QSSSTALS 443
QY 162 NGTVHGDFTSGSSGASQSGRDSHSSSLSEQYPDWPAEDMFD--HPTCELIKGKTKS 220
DB 444 NGFYHFGSTS--SSPISPASSDL-SVAGSLPDTFND--ASPSFGLHPSVPVHVC-----T 493
QY 221 EESL-SDLTGLSLSLQLDLGPSSLDEVLNVMKKNK 254
DB 494 EESLMSLNGSVPSSELDGLDSEKDKML--VEKOK 526

RESULT 4
US-09-957-641-2

; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00
; CURRENT APPLICATION NUMBER: US/09/957,641
; CURRENT FILING DATE: 2001-09-16
; PRIOR APPLICATION NUMBER: US 60/234047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-957-641-2

Query Match 7.9%; Score 105; DB 9; Length 2332;
Best Local Similarity 23.0%; Pred. No. 1.3;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPYLKANNKKGKFKLRDILSPDMIS-----PPLGDFRHTIHIGEGQHDVFGDISF- 58
DB 799 TPHGLSLDLOEAKYETESDPSGAIDSNLSSEMTHERPOLH-----HS--GDMVFT 850
QY 59 -----LQNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVETETPSPVL---KNAIS 108
DB 851 PESGIQLRLNEKL--GTTAATELKKLD-----FKVSTSNLSTIPSDNLAAGTDNTSS 903
QY 109 LPTIGSQALMLPLSPVTFNSKQES--FGPAKLPRLSCEPVMEKAQEKSSLLENGTV 165
DB 904 LG-----PPSMFVHYDSQDITLFGKSSPLTESGGPLSLSENNDSKLLSGLM 953
QY 166 HQGDTWSSGSSASQSS---QGRDSSHSSLSSEQYPDWPAEDMFDPCELIKGTKSEE 222
DB 954 NSQESSWGKNVSTESGRLEKRAHGPALL-----TKDNALFKVSISLKTNTKTSNN 1006
QY 223 SLSD-----LTGSLSLQDLGPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLLIENSPSVWONILSDTEFKKVTPLIHRLMDKN 1055

RESULT 5

; US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601

; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: <unknown>
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-10-187-319-2

Query Match 7.9%; Score 105; DB 9; Length 2332;
Best Local Similarity 23.0%; Pred. No. 1.3;
Matches 67; Conservative 46; Mismatches 102; Indels 76; Gaps 15;

QY 5 TPYLKANNKKGKFKLRDILSPDMIS-----PPLGDFRHTIHIGEGQHDVFGDISF- 58
DB 799 TPHGLSLDLOEAKYETESDPSGAIDSNLSSEMTHERPOLH-----HS--GDMVFT 850
QY 59 -----LQNYELLPGNOEKAHLGQFPGHNEFFRANSTSDSVETETPSPVL---KNAIS 108
DB 851 PESGIQLRLNEKL--GTTAATELKKLD-----FKVSTSNLSTIPSDNLAAGTDNTSS 903
QY 109 LPTIGSQALMLPLSPVTFNSKQES--FGPAKLPRLSCEPVMEKAQEKSSLLENGTV 165
DB 904 LG-----PPSMFVHYDSQDITLFGKSSPLTESGGPLSLSENNDSKLLSGLM 953
QY 166 HQGDTWSSGSSASQSS---QGRDSSHSSLSSEQYPDWPAEDMFDPCELIKGTKSEE 222
DB 954 NSQESSWGKNVSTESGRLEKRAHGPALL-----TKDNALFKVSISLKTNTKTSNN 1006
QY 223 SLSD-----LTGSLSLQDLGPSLDEVLN-----VMDKN 253
DB 1007 SATNRKTHIDGP--SLLIENSPSVWONILSDTEFKKVTPLIHRLMDKN 1055

RESULT 6

; US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; FILE REFERENCE: 6627-PA1170
; CURRENT APPLICATION NUMBER: US/10/132,829
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens

APPLICANT: PLOWMAN, GREGORY
APPLICANT: MARTINEZ, RICARDO
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
FILE REFERENCE: 240/300
CURRENT APPLICATION NUMBER: US/09/291,417A
EARLIER FILING DATE: 1999-04-13
EARLIER APPLICATION NUMBER: US 60/081,784
NUMBER OF SEQ ID NOS: 147
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 13
LENGTH: 1239
TYPE: PRT
ORGANISM: Mammalian (Human) ZC1
US-09-291-417-13

Query Match
Best Local Similarity 26.0%; Score 88.5; DB 9; Length 1239;
Matches 52; Conservative 24; Mismatches 79; Indels 45; Gaps 8;

QY 68 GNOEKALGPPGHNFFRANSTSDSVETETPSVLAISLPTIGSQALMLPLSPVT 127
DB 687 GSQPSHPGSGSGGERFRVRSSSKS--EGSPSQRLENVAKKP----- 727
QY 128 FNSKQESFGPAKLPLSCPEVMEKAQE-----KSSLENGTVHQDTSWSSG 176
DB 728 -EDKKEVFRPLKPADLTA-LAKELRAVEDVRPHKVTDYSSSSSESGTTDEDDVEQEG 785
QY 177 SASQSGGRDSH-SSSLSEQYPDWPA-----EDMFDPH--TPCE---LIKKTKEES 223
DB 786 ADESTSGPEDTRAASSLNLNGETESVKTMTIVHDDVESEPAKTPSKETLIVRTQSASS 845
QY 224 LSDLTGSLSLQLDLPSSL 243
DB 846 TLQKHKSSSSFTPIPRLL 865

RESULT 10

US-08-781-986A-5249
Sequence 5249, Application US/08781986A
Publication No. US20030054436A1
GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5235
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5249:

SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match
Best Local Similarity 21.7%; Score 88; DB 1; Length 936;
Matches 57; Conservative 25; Mismatches 121; Indels 60; Gaps 6;

QY 15 KKGKFKRLDILSPDMISPLGDFRHTHIGKGQHDVFGDISFLQGNVELLPNGQEKH 74
DB 380 KYGKFYNLSIKGTIDQIDKTNTYROTITYNPSGDNYI--APVLTGN--LKPNTDSNAL 434
QY 75 LGQ-----EPGHNEFFRANSTSDSVETE 97
DB 435 IDQNTSIKYKVDNADLSESYFVNPNFEDVTNSVNITFPNPNQYKVEFNTPDQIT- 493
QY 98 TSPVLKNAISLPTIGSQALMLPL-----SPVTFNSKQESFGPAKLRL 143
DB 494 TPYIVVNGHIDPNSKGDALRSTLYGNSNITWRSMWMDNEVAFNNGSGSGDIDKPVY 553
QY 144 SCEPVMEKAQKESLLENGTVHQDTSWSSGSAQSSQGRDSHSSSLSEQYPDWPAED 203
DB 554 --PEQPEDEGEIEPIPEDSDSDPGSDSGSDSNSDSGSDSGSDSTSDSGSDSASDSAS 610
QY 204 MFDPHPTCELIRKKTKEESLSD 226
DB 611 DSDSASDSASDSASDSASDSASDS 633

RESULT 11

US-09-533-029-110

Sequence 110, Application US/09533029
Publication No. US20030046723A1
GENERAL INFORMATION:

APPLICANT: Heard, Jacqueline
APPLICANT: Brown, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Keddle, James
APPLICANT: Pineda, Omatra
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 110
LENGTH: 371
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G1650
US-09-533-029-110

Query Match

Best Local Similarity 6.6%; Score 87; DB 9; Length 371;
Matches 63; Conservative 30; Mismatches 94; Indels 52; Gaps 11;

QY 33 PPLGDF-----RHTHIGKEGQHDVFGDISFLQGNVELLPNGQEKALGQFPCHNEFFRA 87
DB 33 PPLGDF-----RHTHIGKEGQHDVFGDISFLQGNVELLPNGQEKALGQFPCHNEFFRA 87


```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38321
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL023279.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
; OTHER INFORMATION: SWISSPROT HIT: P43363, EVALU 2.00e-48
; OTHER INFORMATION: EST_HUMAN HIT: BE895761.1, EVALU 4.00e-71
US-09-864-761-38321
```

Query Match 6.5%; Score 86; DB 10; Length 473;
Best Local Similarity 23.9%; Pred. No. 8;
Matches 62; Conservative 24; Mismatches 73; Indels 100; Gaps 13;

```
QY 32 SPPGDFRHTIHK---EGQHD-----VFQDISFLQGN 62
Db 34 SPEGPVQSPHSPQSPPEGMHSQSPLOSPEAPEGEDSLPQIPQSPLEGEDSLSLH 93
QY 63 YELPGNOEKA---HGGEPGHNFFR-----ANSTDSV---FTEPSPVLKN 105
Db 94 FPQSPPEWEDSLPLHFPQFPQGEDFQSSLSQSPVSCSSSTSLPQSPFPESPQS---- 149
QY 106 AISLPTIGSQALMLPLSPVT-----FNSKQES-----FGPAKLRLSCEPYME 150
Db 150 -----PPEGPAQS---PLQRPVSSFYSYTLASLQSSHESQSPPEGPAQSPLOS--PVSS 200
QY 151 EKAQEKSSLENGTVHQDTSWGSSGASQSSQGRDSSSSLSSEQYPPWPAEDMFDHPTP 210
Db 201 FPSSTSSLSQSSPV-----SSPPSSTSSLSKSSPESPLQ-----SP 238
QY 211 CELIKGKT-----SEESLS 225
Db 239 VISFSSSTSLSPFSESSS 257
```

RESULT 15
US-09-814-550-2
; Sequence 2, Application US/09814550
; Patent No. US20020102641A1
; GENERAL INFORMATION:
; APPLICANT: Schlavi, Susan
; APPLICANT: Madden, Stephen
; APPLICANT: Manavalan, Parthasarathy
; APPLICANT: Levine, Michael
; APPLICANT: Jan de Beur, Suzanne
; TITLE OF INVENTION: ONCOGENIC OSTEOMALACIA-RELATED GENE 1
; FILE REFERENCE: 5014US
; CURRENT APPLICATION NUMBER: US/09/814,550
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,786
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/241,598
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 525

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-814-550-2
```

Query Match 6.5%; Score 86; DB 10; Length 525;
Best Local Similarity 22.8%; Pred. No. 9.2;
Matches 44; Conservative 26; Mismatches 73; Indels 50; Gaps 8;

```
QY 47 EGQHDVFQDISFLQGNFELPQNG-----EKAHLGOF-----PGHNEFFRANGSTDS 93
Db 336 EGSNDIMGST-----NFKELPGREGNRVDAGSQNAHQKVEFHYPPAPSKERKESGSDA 390
QY 94 V-----FTEP-----SPVLKNAISLPTIGSQALMLPLS-PYTFNSK 131
Db 391 AESTNNEIPKNGKSTRKGYVDHSNBNQATLNKQRFPSKGSQGLPIPSRGLDNEIKNE 450
QY 132 QESF-GPAKLRLSCEPYMEKQEKSSLENGTVHQDTSWG-----SSGASQS 181
Db 451 MDSFNGPSHENIITHGRGYHYVPHQONNSTRNKGMPOGKSGWGRQPHSNRRFSSRRRDS 510
QY 182 SQGRDSSSSLSSE 194
Db 511 SESSDSGSSSES 523
```

Search completed: April 23, 2003, 11:04:32
Job time : 19.6148 secs

